Preliminary Report
Letter Report
NRL [redacted] 7930-230

A SCORING PROCEDURE FOR THE MULTIPLE TARGET CORRELATION AND TRACKING PROBLEM

IRWIN R. GOODMAN

Systems Research Branch
Space Systems Division

557808
(1)

118p

AUGUST 1980

NAVAL RESEARCH LABORATORY
Washington, D. C., 20375

Abstract

This report develops a scoring or, equivalently, validation procedure for those multiple target correlator-trackers which, in effect, form a partitioning $\overset{o}{Q}$ of accumulated data Z from multiple sensor sources into disjoint track sets and a possible false alarm set. The accumulated data may involve missed detections, false alarms, and may describe, typically, measured target positions, velocities, or various target attributes, such as hull length or even identification. One such score, J, for a given correlator-tracker is defined to be $J(\overset{o}{Q},Z) = -2 \log pr(\overset{o}{Q}|Z)$, a form related to the posterior distribution function of partitionings of data. Use of J as a measure can be justified from both information and decision theoretic viewpoints. In particular, minimal J is achieved by the posterior maximum likelihood estimator of Q. $E(\overset{o}{Q},Z)|\overset{oo}{Q})$, a measure of cross entropy, is minimized among constant partitionings by $\overset{o}{Q} = \overset{oo}{Q}$. A closely related score $J'(\overset{o}{Q},Z) = -2 \log pr(Z|\overset{o}{Q})$ has similar properties to J and is computationally more convenient. The emphasis of this report is on J'.

When a linear Gauss-Markov tracking and observation model is assumed, and false alarms and attribute information are modelled by Gaussian processes, J' is shown to be a sum of relatively simple computable terms. The distribution of $(J'(\overset{o}{Q},Z)|\overset{o}{Q})$ under the above assumptions is that of a chi-square random variable plus a constant. Thus, J' is a natural measure of the goodness of fit of a correlator-tracker's output to the data it is operating on.

Several correlator-trackers can be ordered with respect to overall relative accuracy through use of their scores in a weighted sense - depending on the prior decision probabilities and the decision costs involved. However, it is also shown that type I and type II decision errors are difficult to compute, since distributions involving $((J'(\overset{o}{Q},Z) - J'(\overset{oo}{Q},Z))|\overset{o}{Q})$ apparently cannot be obtained in simple form, in general.

CONTENTS

INTRODUCTION

This report is the first part of a two part study concerning performance measures and scoring procedures for the multiple target, multiple sensor correlation problem. In this document, a methodology is developed for determining how well the outputs - in the form of partitionings of the data - of a given correlator-tracker match the actual track sets and false alarms that are present. In the second part of the study, numerical examples will be presented for simulated real-world situations, including ocean surface and subsurface scenarios. These examples should both illustrate the potential use of the scoring procedures developed here, and lead to various sensitivity analyses. In the latter, trade-offs are expected to be established between a number of quantities of interest, including, background shipping densities; types of evasive target motions and group formations; number of (true) targets of interest; false alarm rates; incoming sensor measurement rates; averaged measurement errors; and correlation-tracker design - illustrated, e.g., by the number of, or average length of, track sets formed.

The study is part of the on-going Correlation Handbook Project at the Naval Research Laboratory (Code 7932), under the aegis of NAVELEX. The results of this report extend and supersede much of the earlier work presented in Ref. [1].

A performance (or validity) measure of a given correlator-tracker determines - through its numerical outcomes - how well the outputs of that correlator-tracker estimate the true target - false alarm situation, through its partitioning of the observed data into track sets.

The purpose of this report is to present a scalar performance measure which: is applicable to a wide class of correlator-trackers; is relatively simple in form, suitable for real world operational implementations for use in ocean surface, subsurface, air, or ground surveillance, or in any combination of these scenarios; and is justified by a mathematical-logical basis.

(For additional explanation of terms used here, see the following section.)

The correlator-trackers treated here operate on data, received by possibly several sensor systems, which is in the form of reports that may, unknown to the observer, represent targets of interest or false alarms. The class of sensor systems considered is general and is essentially only limited by the restriction that the time and sensor source of each report be known. Reports may contain: (1) geolocation target or false alarm measurements, such as observed positions or velocities, which typically come from bearing, range, and related sensors; and/or (2) possible non-geolocation target attribute information, such as observations of hull length, number of sensors on-board, emitted frequency information, which usually arise from intelligence sources, visual sightings, acoustic sensors, etc., in conjuction with (1).

One candidate performance measure is J, the negative log posterior distribution of data partionings evaluated at that partitioning characterizing the correlator-tracker, at hand. A closely related measure is J', the negative log conditional distribution of the observed data, conditioned on the partitioning. J' is shown to have a much simpler form than J. If no prior knowledge is available conerning what the true partioning of data is, then a uniform distribution assumption lends to J and J' differing by a function of the data only - and not of any partitionings. In this case, all of the theoretical justifications for J equally apply to J'. The bulk of this report concerns the evaluation and presentation of the properties of J'.

J' can be expressed as the sum of three submeasures; Evaluation of J' involves computing Kalman filters for the various track sets of the correlator-tracker, resulting in a submeasure of goodness of fit of the tracks. Another submeasure computes the plausibility of the false alarm set determined by the correlator-tracker. A third submeasure determines the goodness of fit of the target attribute data at hand (such as observed ship identifications, frequencies received, hull lengths observed, etc.). These submeasures, from their relatively simple structures, are expected to be implementable for real time incoming data, recursively in many instances.

In order to properly determine J', Kalman filter outputs of the correlator-tracker being evaluated may have to be disregarded and replaced by the specially structured Kalman filters used for J'. This, in general, is due to either or both of the following two factors being present:

1) The correlator-tracker at hand may in its own operational form be using an adaptive extended - or in some way non linear-Gauss-Markov target motion and measurement models. In this case, present or relatively recent target positions, e.g., will be more accurately estimated than previous ones, and correlation errors can actually be worse than for non-adaptive filters. In any case, sums of the apparent quadratic forms of data fit do not really represent the fit of the entire track set to the data.

2) Targets may be maneuvering so that the linear-Gauss-Markov target motion assumptions become invalid.

Appendix C exhibits detailed flow charts for computations required of incoming data in determining J', with some particular numerical examples given in Appendix D.

BACKGROUND

In brief, the correlation problem considered here may be defined as follows: Determine, within feasible computational bounds, that partitioning of the observed data set, which most accurately reflects through its component track sets, the true (but unknown to the observer) partitioning. In the correlation problem considered here, data is accumulated in the form of vector valued reports, labeled as to sensor system source and time, accumulated over some specified sampling interval. The component sets of the partitioning represent either track sets, each corresponding to a single target source (estimated, if the partitioning itself is an estimate of the true partitioning), or the false alarm (or clutter) set. (The latter may be vacuous.) The observed data vectors may be classified into two categories: geolocational and non-geolocational attributes. For any given time and sensor system it is possible that either of these types of data vectors may be vacuous. For the geolocational case this may be due to missed detections; however, generally when geolocation information is not missing, the number of components in the data vector will be fixed and typically includes positional and/or velocity observations. False alarms may be mixed into this data. For the non-geolocation case, the number of components and type of information available in a given data vector is random. Typically, non-geolocation target attributes can include any subset of the following: sighting of ship identification, observed ship silhouette/shape, estimated hull length, observed flag color, number of sensors of a given type detected to be on-board vessel, and ship machinery frequency outputs observed.

Broadly speaking, correlator-trackers, i.e., schemes which first establish partitions of the data and then use these, in turn, to establish tracks of the targets, may be divided into two classes: 'soft decision' and 'hard decision' types. The former, basically establish posterior distributions or sets of possible partitions to be considered, generate weights based on the likelihoods or similar measures of the possible partitions, and then estimate target motion parameters through weighted sums of Kalman filters. (See, e.g., Ref. [2]). On the other hand, hard decision correlator-trackers may be identified with a single partitioning of the data; the estimated target motion parameters being determined by a Kalman filter for each component track set of the partitioning. (See Ref. [3].)

In this report, only those correlator-tracker schemes are considered which can be identified with a single partitioning of the accumulated data, for each sample time. Thus, this includes all hard decision correlator-trackers and those soft decision ones which can be suitably modified, such as by choosing that partitioning of the data which maximizes the posterior distribution of possible partitionings, or some related statistic.

Extensive overviews and general analyses of the multiple target multiple sensor correlation problem may be found in NRL's Correlation Handbook [4]; see also the earlier work of Kullback and Owens [5]. General models for the problem are presented in Sittler [6], Reid [2] and Goodman [7]. Bar Shalom [8] and Goodman [7] Chapter 12 also contain technically detailed surveys of various correlator-tracker schemes, with a number of examples presented for both the hard and soft decision cases.

STATEMENT OF THE PROBLEM - PERFORMANCE MEASURES

Consider the problem of determining how well a particular correlator-tracker technique works. It is obvious that if target density or false alarm rates are sufficiently high, even with very accurate sensing systems present, no correlator-tracker will perform well. Other factors also contribute to correlation difficulties, such as sensor reliability, data rates and variable types of target motion, including maneuvers. Despite the above negative remarks, there appear to be many occasions (moderate data rate with few targets and false alarms present, e.g.) when the correlation problem may be amenable to reasonably accurate solutions. In these situations, an analytically based performance measure can serve as a valuable tool for ranking and comparing quantitatively different approaches.

Due to the possible combinations of many factors - including target initialization and termination times within the sensor areas of interest, detection failures, variable target motions and maneuverings and matching of non-geolocation attribute information to the correct track sets - a large number of possible partitionings of the data into track sets can occur. This, in turn, can lead to an exponential growth - with respect to sampling time - of the number of computations required in implementing or monitoring error bound performance of correlator-trackers, even under very simplistic model assumptions. (See, e.g. Theorem 4 of the abridged version of Ref. [7] for an example of exponential growth in the number computations needed to compute target motion parameters. See also Ref. [1].) Thus, alternative performance measures must be sought which do not involve such a large amount of computations.

Performance measures, or, as they are often called MOE's (measures of effectiveness) can be roughly divided into two basic classes: (1) observer dependent only and (2) simulator-observer dependent.

A performance measure is defined to be observer-dependent only, if it is a true statistic, i.e., known function, of the cumulative data and the partitioning determined by the correlator-tracker in question.

On the other hand, a performance measure may depend in implementation on knowledge of what data constitutes the true track sets and/or false alarms, i.e., on the true partitioning of the data. The measure may further depend on knowledge of the correct target motion parameter vectors such as positions, and correct time of initialization and termination of the targets from the region being surveilled. In this case, the performance measure is said to be simulator-observer dependent.

In general, an observer dependent only performance measure, or briefly, score, is evaluated relative to its statistical distribution as a random quantity (based on the random data), or at least relative to its statistical moments, conditioned on an assumed true situation (i.e., true partitioning of data or even true target motion parameters). These statistical moments, are thus simulator-observer performance measures generated by the score.

Some examples of scores are: average observed track depth or persistance; number of tracks exceeding a depth of three sample times; and number of tentative tracks, i.e., track sets having no more than a prescribed upper bound on track depth (e.g., two). (See, e.g., References [9] and [10].) Another measure that is observer dependent only is consistancy of performance (see [10]). The latter measure describes how much difference exists between the initial filtered-partitioned output data from a given correlator-tracker and a second output resulting from recycling the initial output through the same correlator-tracker (by using it formally as raw input data). (See also Ref. [1] pages 6 and 7 for further mathematical details.)

Simulator-observer dependent measures of performance are much more plentiful in the literature of correlator-trackers and related algorithms than those in the former class. Wiener et al. [11] present a number of examples of these measures, as well as a general overview of the rule MOE's in general play in the command and control aspects of ocean surveillance. (See also the overviews in [12] and [13].) Typical measures here include all measures of accuracy of target motion parameter estimators, thus, e.g., estimated coverage probabilities of target parameters; track purity, i.e., percentage of correct elements in any given track set; probability of correct association between new reports and previous established tracks, and percentage of targets having track depths of given lengths. Most of the above measures may be averaged with respect to sampling times and number of targets present.

This section is concluded with a brief overview of the literature of simulator-observer dependent MOE's.

Among the measures Willman [9] uses for evaluating the performance of his hard decision 'matrix scan' correlator-tracker are four involving actual and estimated track depths (the latter obtained through estimated probabilities of consecutive track linkages). In addition, Willman employs: a measure which is formally the same as the determinant of a corresponding two-by-two table of entries;

percentage of old targets having established tracks lacking new assignment of reports; and percentage of misassigned reports.

Kullback and Owens ([5], pp. 31-42) consider numerical differences between the number of targets and false alarms estimated and actually present, as well as track purity. In addition, use of non-geolocation attributes by a correlator-tracker is proposed to be measured by an ambiguity function which counts the total number of possible ships in the area which could fit those attributes assigned to each track set. Other measures are discussed which measure the information level of correlator-trackers.

Adler et al. [14] have proposed a number of interesting MOE's, including the rate of track fragments (interruptions in tracking the same targets) and average time between correct correlations (or linkages) of old tracks to newly arriving reports. Turner and Marder [15], in addition to the usual measures, use as a measure the number of observations required of a target to form a firm track. Reid [2] proposes (in addition to percentage of targets tracked, lost, etc.) the average number of partitionings of data kept, following pruning and merging, according to (his soft decision) correlator-tracker. (See also the MOE's presented in References [16] - [18].) The performances of the large scale correlator-tracker presented in [19] (resulting in various model improvements - Personal Correspondence) were measured by a number of MOE's previously mentioned, including track purity, accuracies of estimators, and the number of tracks sets having relatively short depths.

Observer dependent measures are presented in greater detail in the next section.

OBSERVER DEPENDENT PERFORMANCE MEASURES - SCORES

As mentioned in the previous section, many measures of performance exist, but in effect they form a patchwork quilt of descriptions for the behavior of correlator-trackers in general.

Alspach and Lobbia [20] construct an observer dependent only performance measure, which operates on a correlator-tracker's perceived data partitioning only through the total number of reports assigned to track sets and the total number of reports assigned to the false alarm set. The statistical expectation of the evaluation of this measure (with respect to the randomness of the data) holding the number of track reports and false alarms fixed, is also a performance measure which is both observer and simulator dependent. This is shown to be a concave unimodal function (under certain simplifying assumptions) of the total number of reports assigned to track sets. This function is seen to be a linear function when the number of reports assigned to track sets is less than or equal to the true number of reports that should be assigned (via the true partitioning) to the track sets, and possesses an absolute minimum when the number of reports decided to be targets is equal to the number of actual targets.

In this report a score J' is proposed which is expected to be relatively feasible to implement: (twice) the negative logarithm of the conditional probability function of the possible data partitionings, evaluated at that partitioning determining the correlator-tracker to be evaluated. An associated simulator-observer dependent performance measure (I') is also presented. This measure is essentially the statistical expectation of J' conditioned on any given true partitioning of data, a cross entropy measure.

Although Alspach and Lobbia [20] briefly discuss the possible use of a statistic related to the one proposed here (but do not consider non-geolocation attributes), they dismiss use of it because of apparent difficulties in determining the non-random terms consisting of determinants of innovation covariance matrices. In effect, in Alspach and Lobbia's score, the non-random terms are replaced by the product of the number of data points decided to be clutter or false alarms employing undetermined constant scores for any single point decided to be clutter.

In Reference [21] and [3] (the former treating data recursively in time, the latter handling data non-recursively), new reports are assigned to those track sets and to the false alarm set such that performance measures which are modifications of the log likelihood of possible partitions - a statistic related in form to that proposed in this report - are maximized (or minimized). (See Ref. [8] for concise descriptions of other correlation techniques which, in effect, also use performance measures for determining implementations.)

The motivation of the choice of the negative log conditional or posterior probability functions of possible partitionings as a measure of a correlator-tracker's performance is based on five desirable properties:

1. The measure in its initial form (before taking expectations

with respect to the data) is truly observer-dependent only, and may thus be used as a real world scoring method.

2. The score is relatively simple in form, or can be reasonably approximated by a simple structure, suitable for real-time implementations.

3. The score directly reflects the goodness-of-fit of the partitioning of the correlator-tracker in question to the given data, employing geolocation terms, non-geolocation attributes and false alarm data.

4. The statistical distribution of the score is related to a chi-square random variable.

5. Use of the score can be justified from information theory and statistical decision theory viewpoints.

The statistical expectation of the score is observer-simulator dependent and can also be used as a measure of performance. In particular, the expectation of the measure will have smaller values, generally, for correlator-trackers which use more information and/or have lower risks (equivalently, better approximate the Bayes or maximal posterior data partitioning).

The above five properties may well serve as a general guideline for establishing performance measures for correlator-trackers.

OUTLINE OF THE REPORT

The introductory sections describe the correlation problem, in general, and performance measures, in particular. 'Hard' vs. 'soft' types of correlator-trackers are detailed; the former being more conducive for being evaluated by the performance measures J and J' presented in this report. J is essentially the negative log posterior distribution function evaluated at that partitioning Q of data characterizing the given correlator-tracker, while J' is a related more computationally simpler measure involving the conditional distribution of the data (conditioned on Q). Surveys of the available literature for both observer dependent (J and J' are in this class) and simulator-dependent performance measures are presented. Relations between these measures and the ones proposed in this report, where applicable, are described. A guideline is presented for desirable properties that any performance measure for correlator-trackers should possess. (J' satisfies these criteria.)

The Analysis Section first establishes (subsection 2) comprehensive definitions and mathematical models for all assumed target motions (eq. (2.1)), occurrence and location of false alarms (subsections 2((3)), ((6))), and observation measurements (eq. (2.2)). Following this, statistical decision theory and information theory bases are established for use of the proposed scores (subsection 3). In essence, it is shown that the lower the value of J or J', the closer the correlator-tracker (through its partitioning of the data) matches the true but unknown partitioning of the data into the correct track sets and false alarm set.

The next subsection (4) of the Analysis Section develops the full structure of the simpler performance measure J'. This leads to a decomposition of J' into a sum of three terms, reflecting goodness-of-data-fit to the partitioning determined by a given correlator-tracker (eq. (4.1)). The geolocation target data term is shown to be the sum of constant and quadratic forms of innovations that are outputs from Kalman filters (eqs. (4.2) - (4.51)). The (geolocation) false alarm term is the sum of constant and quadratic forms of data (eq. (4.52) - (4.61)). Both terms are seen to depend explicitly on the number of targets and false alarms perceived by the correlator-tracker. The non-geolocation target attribute term involves discrete sums of probabilities (eqs. (4.62) - (4.70)). If a normal approximation is made in the modeling of these attributes (see remarks following eq. (4.70)) then a Kalman filter (eqs. (4.114) - (4.127)) may be used for evaluation of this term: a constant plus quadratic forms of the data. (See eqs. (4.63), (4.64), (4.68), (4.69), (4.79) - (4.82).) A further simplifying approximation - which avoids basically prior knowledge of the randomness of the non-geolocation characteristics is given in Appendix B; in addition the accuracy of this approximation is also demonstrated.

The final subsection (5) of the Analysis Section develops the distributional properties of J'. It is shown that J', conditioned on the partitioning determining the correlator-tracker being evaluated, is distributed as the statistical independent sum of a constant, a chi-squared random variable, and a discrete random variable (see eqs. (5.1) - (5.5)). Under normal distributional approximations for the non-geolocation attributes, the above discrete random variable is replaced by the sum of a constant and a chi-squared random variable (see remarks following eq. (5.5).) It follows in this case that J' is distributed as the sum of a constant and a chi-squared random variable (eqs. (5.6) - (5.9)). It is also shown that J' conditioned on a partitioning of data not coinciding with the one determining the correlator-tracker being evaluated, has a distribution which is not easily computable (not even a non-central chi-square distribution). (See equations and remarks following eq. (5.9).)

Appendix A presents a procedure for calculating matrix inverses in prescribed block form. This can be useful in evaluating parts of the Kalman filters used in the computations for the geolocation targets and non-geolocation attributes terms (assuming a normal approximation).

A further approximation to a quadratic form arising from normal distributional approximations for the non-geolocation attributes is presented in Appendix B, where also error bounds are derived. Use of this approximation minimizes required prior knowledge of the randomness of the true non-geolocation attributes of the targets.

A complete set of flow charts for computing J' relative to incoming data is presented in Appendix C.

ANALYSIS

1. INTRODUCTION

In order to be able to define and analyze consistantly the performance measures proposed here, a rigorous mathematical - logical model is established for the general correlation-tracking problem. (This largely simplifies the model proposed in Ref. [7].) This model consists of eight key aspects: ((1)) sensor systems, ((2)) target initializations and terminations, ((3)) existance of target state parameter vectors and detection, ((4)) partitionings of observed data, ((5)) target state parameter vectors, ((6)) observed geolocation data, ((7)) non-geolocation attribute data, and ((8)) total observed data.

Although the model obviously simplifies the real-world situation, it is expected that its constituent assumptions are 'reasonable' approximations of reality. In keeping with the attempt to be as faithful as possible to the real world setting, the model established represents a relatively short data sampling period, j_0, and can be changed with respect to each new sampling period; all of the available output information from the previous sampling period, being used as input - prior information for the next period. This disjointing of the sampling times into short segments should make more valid the homogeneous linear Gauss-Markov motion and measurement models used in ((5)) and ((6)). The latter assumption can account to some degree for variations in types of target motion by allowing in the model a reasonably large state vector dimension with zeros possibly occurring in particular entries. (Thus, quadratic polynomial motion includes as a special case, straight line motion, with the possibility of estimated error covariances being actually larger than necessary.) This results in the avoidance of a potentially large branching problem that can arise in attempting to model the general correlation-tracking problem. (See, e.g. Goodman [7], especially Theorems 4 and 5 of the abridged version for an illustration of the complexity arising when variable target motions and maneuvers are modeled.)

Even over a relatively short sampling period, a target may engage in maneuvering - such as zig-zags or circular motion - which is not really modeled by the same linear Gauss-Markov target motion model. Yet, a given correlator-tracker may still retain the ability to follow that target, and thus essentially put all observations of it into the same track set. (This will usually be carried out by highly nonlinear adaptive Kalman filters - which do not reflect the target motion model assumed here.) Clearly, a linear Gauss-Markov fit to this track set - which is one of the computations required to obtain the score of the correlator-tracker in question (see eq. (4.20)), - in general for this situation is not appropriate.

Consequently, a special procedure is used in computing the score for a correlator-tracker which has at least some track sets generated apparently by a maneuvering target as described above. In essence, the procedure retains the mathematical rigor of the model, by simply replacing the single track set in question by a disjoint union of different (but almost contiguous) track subsets, each distinct track subset based on measurements eminating from a different linear Gauss-Markov motion model (as is permitted in the overall model). The termination time of each such subset (and consequently, the next sampling time being the initial time of the following subset) is determined by either a simple chi-squared test involving goodness of fits or, more generally by using the non-geolocation attribute of common - but unknown - identity in monitoring the entire goodness of fit data probability function.

See Appendix C, for implementation of this procedure and for overall flow charts for computing the score J' of a given correlator-tracker as a function of incoming data.

It should also be noted that often correlator-trackers operate with nonlinear adaptive-extended Kalman filters of the track sets. In these cases the original filters must be replaced by the linear ones for the model developed here (see eqs. (2.1), (2.2)) in order to reflect the total goodness of fit of the track sets.

2. ASSUMPTIONS AND BASIC DEFINITIONS FOR THE CORRELATION-TRACKING MODEL

((1)) Sensor Systems

q known fixed sensor systems sample data in the form of vector reports at known and possibly random outcome times $t_0 < t_1 < \ldots < t_{j_0}$, where j_0 is chosen relatively small based on experience or other factors. (See the discussion on choice of j_0 in the previous subsection.)

((2)) Target Initializations and Terminations in AOI

An unknown number $M^{(j)}$ (to the observer) of targets begin existance in the sensor areas of interest (AOI) up to time t_j. This is due to targets either entering the AOI for the first time and/or becoming sufficiently active that they may be detected by at least one of the sensor systems present. Some of the targets in the AOI, later may terminate existence (relative to the AOI), i.e. exit the AOI or quiet down so that they can no longer be detected by any of the q sensor systems present. The number of targets actually existing at t_j is denoted by M_j, also unknown. Thus, $M_j \leq M^{(j)}$, $M^{(j)}$ is non-decreasing in j, and M_j and $M^{(j)}$ are random integer outcomes.

Related to the above definitions, associated with any target $i \in M^{(\infty)} \overset{df}{=} \bigcup_{j=1}^{\infty} M^{(j)}$ are two unknown random integers u_i, v_i, $0 \leq u_i \leq v_i$, where t_{u_i} is the time of target i's initial existence in the AOI and t_{v_i} is the time of its termination. It is assumed that any target exists over successive sampling times and if it ceases existence (for two or more sampling times) and begins again at a later time, it is considered a distinct target here.

((3)) <u>Existence of Target State Parameter Vectors and of Observed Data Vectors</u> (Detection)

For any target i, time t_j and sensor k, X_{ij} represents its state parameter vector (not dependent on k), Z_{ijk} its observed geolocation measurement vector, and Y_{ijk} its observed non-geolocation attribute measurement vector. Z_{0jk} is the set of false alarm vectors occurring at time t_j due to sensor k.

$X_{ij} \neq \phi$ iff target i exists at time t_j, in which case $\dim(X_{ij}) \equiv m$ is known. Thus $X_{ij} \neq \phi$ iff $u_i \leq j \leq v_i$

The following is closely connected with sensor detection:

$Z_{ijk} \neq \phi$ iff sensor k makes at time t_j one geolocation measurement of target i, in which case $\dim(Z_{ijk}) \equiv r_{jk}$ is known.

$Z_{0jk} \neq \phi$ iff sensor k receives at time t_j, f_{jk} $(f_{jk} \geq 1)$ geolocation measurements of false alarms in which case, the false alarm set is $Z_{0jk} \overset{df}{=} \{Z_{0jk\omega} | \omega = 1,2,\ldots, f_{jk}\}$, where $Z_{0jk\omega}$ is the ω^{th} false alarm vector seen by sensor k at time t_j and $\dim(Z_{0jk\omega}) \equiv r_{jk}$ is known. f_{jk} is an unknown random outcome, distributed exponentially as Expo (λ_{jk}), λ_{jk} known. f_{jk}'s are statistically independent with respect to different (j,k)'s.

$Y_{ijk} \neq \phi$ iff sensor k makes at time t_j one non-geolocation target attribute measurement of target i, in which case $1 \leq \dim(Y_{ijk}) \leq b$, $\dim(Y_{ijk})$ and b are known.

X_{ij} is always unobserved.

Z_{ijk} and Y_{ijk} are always observed.

Index $i \geqq 0$ (including $i=0$ for the Z_{0jk}'s) is unknown to the observer, while j and k are always known. Index ω is always unknown. However, when the observer forms a partitioning $Q^{(j)}$ of the data (see the next subsection), relative to $Q^{(j)}$ assumed to be formally true, index i (including $i=0$) becomes (formally) known.

If $X_{ij} = \phi$, then clearly $Z_{ijk} = Y_{ijk} = \phi$, for $1 \leqq k \leqq q$.

On the other hand, $X_{ij} \neq \phi$ does not guarantee Z_{ijk} and/or Y_{ijk} being nonvacuous. For example, sensor k may miss a distinct detection due to resolution problems, noisy background, or reliability problems caused by equipment or human operator failure.

Define, for any i; $1 \leqq i \leqq M^{(j)}$; j, k,

$$\mathcal{Z}_{ijk} \stackrel{\text{df}}{=} \begin{pmatrix} Z_{ijk} \\ Y_{ijk} \end{pmatrix}; \quad \mathcal{Z}_{0jk} \stackrel{\text{df}}{=} Z_{0jk}$$

$\mathcal{Z}_{ijk}$ represents all observed data of target i (for $i \geqq 1$) at time t_j by sensor system k. It is called a <u>report</u>.

((4)) <u>Partitionings of Observed Data</u>

For any i, $1 \leqq i \leqq M^{(j)}$, define

$$\begin{aligned} Q_i^{(j)} \stackrel{\text{df}}{=} \{(i, \alpha, k) \mid & \text{ for all } \alpha, k, \\ & 0 \leqq \alpha \leqq j \ \& \ 1 \leqq k \leqq q \text{ such that} \\ & \mathcal{Z}_{i\alpha k} \neq \phi\} \\ = & \text{ target track index set } i \text{ up to } t_j. \end{aligned}$$

$$Q_0^{(j)} \stackrel{\text{df}}{=} \{(0,\alpha,k,\omega) \mid \text{for all } \alpha, k, \omega,$$
$$0 \leq \alpha \leq j \;\&\; 1 \leq k \leq q \;\&\; 1 \leq \omega \leq f_{\alpha k} \;\&\; \text{such that}$$
$$Z_{0\alpha k} \neq \phi \;\; (\text{and hence } f_{\alpha k} \geq 1)\}$$
$$= \text{ false alarm index, set up to } t_j \text{ (possibly vacuous).}$$

The partitioning index of data up to t_j is given by

$$Q^{(j)} \stackrel{\text{df}}{=} \{Q_i^{(j)} \mid i \in A^{(j)}\},$$

where

$$A^{(j)} \stackrel{\text{df}}{=} \{i \mid 0 \leq i \leq M^{(j)} \;\&\; Q_i^{(j)} \neq \phi\}$$

corresponds to the set of all distinct target track sets established up to t_j, including the false alarm set as a special track set (0).

Notice that $Q^{(j)}$ is determined by the (detection) set of all (i,α,k)'s for which $Z_{i\alpha k} \neq \phi$, and indirectly by all T_{ikj}'s, $M^{(j)}$, by all u_i's, v_i's; and by the set of all $(0,\alpha,k,\omega)$'s for which $f_{\alpha k} \geq 1$.

Given index $Q^{(j)}$, there are infinitely many corresponding partitioning outcomes $Q^{(j)}(Z^{(j)})$, depending on the values of the $Z_{i\alpha k}$'s, where

$$Q^{(j)}(Z^{(j)}) \stackrel{\text{df}}{=} \{Q^{(j)}(Z^{(j)}) \mid i \in A^{(j)}\},$$

$$Q_i^{(j)}(Z^{(j)}) \stackrel{\text{df}}{=} \{Z_{i\;k} \mid (i,\alpha,k) \in Q_i^{(j)}\}$$

In general, we identify, if ambiguity does not arise, $Q^{(j)}(Z^{(j)})$ with $Q^{(j)}$.

Note that all partitionings $Q^{(j)}$ are <u>unlabeled</u>, that is the indices i for each track set (or false alarm set, for $i = 0$) $Q_i^{(j)}$ does <u>not</u> <u>identify</u> what target i really is. It is just a convenient index.

In some of the following equations equality holds only after a suitable rearrangement of indices is made.

((5)) <u>Target State Parameter Vectors</u>

The results here do <u>not</u> depend on $Q^{(j)}$:

Given outcomes (u_i, v_i) for $i = 1,.., M^{(j)}$ and equivalently those (i,j)'s for which $X_{ij} \neq \phi$;

$$X_i^{(j)} \stackrel{df}{=} \{X_{i\alpha} \mid u_i \leq \alpha \leq \min(v_i, j)\}$$

$$= \{X_{i\alpha} \mid 0 \leq \alpha \leq j \ \& \ X_{i\alpha} \neq \phi\}$$

$=$ set of state vectors of target i up to t_j.

$$X_j \stackrel{df}{=} \{X_{ij} \mid 1 \leq i \leq M^{(j)} \ \& \ X_{ij} \neq \phi\}$$

$=$ set of all state vectors of targets existing at t_j.

$$X^{(j)} \stackrel{df}{=} \{X_\alpha \mid 0 \leq \alpha \leq j \ \& \ X_\alpha \neq \phi\}$$

$=$ set of all state vectors of targets existing sometime up to t_j.

Initial state vector X_{i,u_i} of target i (at t_{u_i}) is assumed to be distributed normally $N_m(E(X_{i,u_i}), \mathrm{Cov}(X_{i,u_i}))$, where $\mathrm{Cov}(X_{i,u_i})$ is known and $E(X_{i,u_i})$ is unknown, unless otherwise specified.

(Prior information for X_{ij}, $j < u_i$, may be used to determine both or either of its moments.)

If $u_i < v_i$, for $\alpha = u_{i+1},..,v_i$, assume

$$X_{i,\alpha} = \Phi_\alpha \cdot X_{i,\alpha-1} + G_\alpha \cdot W_{i\alpha}, \tag{2.1}$$

linear Gauss-Markov homogeneous target motion, where

Φ_α is m by m known transition matrix.

G_ℓ is m by n known coefficient matrix.

$W_{i\alpha}$ is distributed $N_n(0, P_\alpha)$; P_α known.

$W_{i\alpha}$'s are statistically independent for different (i,α)'s ; they are all unobserved random variables.

$X_{i,\alpha-1}$ and $W_{i\alpha}$ are statistically independent,

$X_i^{(j)}$'s for different i's are statistically independent.

((6)) <u>Observed Geolocation Data</u>

$$Z_i^{(j)} \overset{\text{df}}{=} \{Z_{i\alpha k} \mid 0 \leq \alpha \leq j \ \& \ 1 \leq k \leq q \ \& \ Z_{i\alpha k} \neq \phi\}, \quad \text{for any } i, \ 0 \leq i \leq M^{(j)},$$

= geolocation data for track set i (and hence for target i, i unknown to the observer) up to time t_j,

$$Z_{ij} \overset{\text{df}}{=} \{Z_{ijk} \mid 1 \leq k \leq q \ \& \ Z_{ijk} \neq \phi\},$$

= geolocation data for the i^{th} track set at t_j (from all sensors),

$$Z_j \overset{\text{df}}{=} \{Z_{ijk} \mid 0 \leq i \leq M^{(j)} \ \& \ 1 \leq k \leq q \ \& \ Z_{ijk} \neq \phi\}$$

= geolocation data for all track sets at t_j,

$$Z^{(j)} \overset{\text{df}}{=} \{Z_{i\alpha k} \mid 0 \leq i \leq M^{(j)} \ \& \ 0 \leq \alpha \leq j \ \& \ 1 \leq k \leq q \ \& \ Z_{i\alpha k} \neq \phi\}$$

$$= \{Z_\alpha \mid 0 \leq \alpha \leq j \ \& \ Z_\alpha \neq \phi\}$$

= geolocation data for all track sets up to t_j

= total observed geolocation data up to time t_j.

Define similarly, $z_i^{(j)}$, z_{ij}, z_j, $z^{(j)}$, etc.

For any $Z_{ijk} \neq \phi$, $i = 1, \ldots, M^{(j)}$, the observed geolocation target data or measurement is assumed to have a linear regression relationship with respect to a corresponding target state vector:

$$Z_{ijk} = B_{jk} \cdot X_{ij} + V_{ijk}$$

In more compact form,

$$Z_{ij} = B_{\underset{i}{j}} X_{ij} + V_{ij} \quad ; \tag{2.2}$$

$$B_{\underset{i}{j}} \stackrel{df}{=} \begin{pmatrix} B_{j1} \\ \cdot \\ \cdot \\ B_{jq} \end{pmatrix}_{\substack{\text{for } 1 \leq k \leq q \\ \text{such that} \\ Z_{ijk} \neq \phi}}, \quad V_{ij} \stackrel{df}{=} \begin{pmatrix} V_{ij1} \\ \cdot \\ \cdot \\ V_{ijq} \end{pmatrix}_{\substack{\text{for } 1 \leq k \leq q \\ \text{such that} \\ Z_{ijk} \neq \phi .}}$$

B_{jk} is always a known r_{jk} by m matrix. V_{ijk} is unobserved and distributed statistically independent of X_{ij} and indeed of $X^{(j)}$; V_{ijk}'s are statistically independent for different (i,j,k)'s.

V_{ijk} is distributed normally as $N_{r_{jk}}(0, R_{jk})$, with R_{jk} always known.

Thus, for $i = 1, \ldots, M^{(j)}$:

V_{ij} is distributed as $N_{s_{ij}}(0, R_{\underset{i}{j}})$,

where

$$R_{\underset{i}{j}} \stackrel{df}{=} \begin{pmatrix} R_{j1} & & 0 \\ & \ddots & \\ 0 & & R_{jq} \end{pmatrix}_{\substack{\text{for } 1 \leq k \leq q \\ \text{such that} \\ Z_{ijk} \neq \phi ,}}$$

$$s_{ij} \stackrel{df}{=} \sum_{\substack{1 \leq k \leq q \\ \text{such that} \\ Z_{ijk} \neq \phi}} r_{jk} = \dim(Z_{ij}) \ .$$

Since r_{jk} is known to the observer, if $Q^{(j)}$ is given, then

s_{ij} will be known (even though i remains unknown to the observer), if $Q^{(j)}$ is given. Similarly for B_{ij} and R_{ij}.

Let

$$Z_+^{(j)} \stackrel{df}{=} \{Z_i^{(j)} \mid 1 \leq i \leq M^{(j)},\ Z_i^{(j)} \neq \phi\}$$

$$= \{Z_{i\ell k} \mid 1 \leq k \leq q,\ 0 \leq \ell \leq j,\ 1 \leq i \leq M^{(j)},\ Z_{i\ell k} \neq \phi\}$$

= total observed geolocational (true) target data up to t_j.

Note that

$Z_0^{(j)}$ = total observed geolocational false alarm data up to t_j.

For $i = 0$ such that $Z_{0jk} \neq \phi$, and for all ω, $1 \leq \omega \leq f_{jk}$, it is assumed that $Z_{0jk\omega}$ is statistically independent with respect to different (j,k,ω)'s, all other $Z_i^{(j)}$'s, $1 \leq i \leq M^{(j)}$, and to $X^{(j)}$. In addition, it is assumed that i $f_{jk} \geq 1$, for fixed j,k, the $Z_{0jk\omega}$'s for $1 \leq \omega \leq f_{jk}$ are identically distributed as $Z_{0jk1} : N(\theta_{jk}, M_{jk})$, where $\theta_{jk} \stackrel{df}{=} E(Z_{0jk1})$ and $M_{jk} \stackrel{df}{=} Cov(Z_{0jk1})$; θ_{jk} and M_{jk} are assumed known. Although in general f_{jk} is unknown to the observer, given (true) $Q^{(j)}$, f_{jk} becomes known. Similar properties are assumed for the $Z_{0jk\omega}$'s given $Q^{(j)}$.

((7)) <u>Non-geolocation Target Attribute Data</u>

Let a and b be fixed known positive integers and $C \stackrel{df}{=} \{C_1, .., C_a\}$, a fixed known set of distinct objects where each $C_\alpha = (C_{\alpha\beta})_{1 \leq \beta \leq b}$; $1 \leq \alpha \leq a$.

The $C_{\alpha\beta}$'s represent possible non-geolocation attributes of any target.

For example for $b = 3$,

$C_{\alpha 1}$ = visual identification

$\varepsilon \{n_1, \ldots, n_{100}\}$,

where each n_i is a distinct ship name,

$C_{\alpha 2}$ = observed flag color

ε {Red, Green, Brown, White} ,

$C_{\alpha 3}$ = no. of radar sensors determined to be on-board the ship

$\varepsilon \{0, 1, 2, .., 10\}$.

Thus, if all possible combinations of attributes can occur, then $a = 100 \cdot 4 \cdot 11 = 4400$; but if certain combinations of values of $C_{\alpha 1}, C_{\alpha 2}, C_{\alpha 3}$ are excluded, then a may be a good deal less than 4400. (This will often be the case.)

Define, random variable H_i (not dependent statistically on $Q^{(j)}$) over the set C to represent the distribution of possible attributes occurring for target i, $i = 1, .., M^{(j)}$. The probability function of H_i is known. (H_i's outcome is unknown to the observer.) Also assume that the distribution function for H_i is <u>not</u> dependent on t_j, for the duration of target i's existence in the A.O.I.

Now for any i,j,k such that $X_{ij} \neq \phi$ and $1 \leq k \leq q$, let T_{ijk} be always an observed random subset of $\{1, .., b\}$ which is statistically independent for different (i,j,k)'s and statistically independent of $X^{(j)}$, $Z^{(j)}$ and $Q^{(j)}$. Let α' and α'' be arbitrary, $1 \leq \alpha', \alpha'' \leq a$ and let $Y' \overset{df}{=} (C_{\alpha', \beta})_{\beta \in T_{ijk}}$ and $Y'' \overset{df}{=} (C_{\alpha'', \beta})_{1 \leq \beta \leq b}$.

Define conditional random variable Y_{ijk} by

$$pr(Y_{ijk} = Y' | H_i = Y^{\sim}, Q^{(j)}) \stackrel{df}{=} p(Y', Y^{\sim}; Q^{(j)}) , \qquad (2.3a)$$

a known function of its arguments (not dependent functionally on i nor k and only dependent on $Q^{(j)}$ through the outcome of T_{ijk}).

Analogous to the definitions in subsections ((5)) and ((6)) define $Y_i^{(j)}$, Y_{ij}, $Y^{(j)}$, $H^{(j)}$, etc.

Y_{ijk} can be considered the observation of H_i, the actual non-geolocation attribute present (given selection outcome T_{ijk}) of target i.

Assume that given the T_{ijk}'s, all of the $(Y_{ijk} | H_i, Q^{(j)})$'s are statistically independent with respect to different (j,k)'s; suppose all $(Y_i^{(j)} \mid H_i, Q^{(j)})$'s are statistically independent for different i's,

Equivalently, for outcome $T_{ijk} \neq \phi$

$$Y_{ijk} = B_{ijk} \cdot H_i + V_{ijk}$$

In more compact form,

$$Y_{ij} = B_{ij} \cdot H_i + V_{ij} ; \qquad (2.3b)$$

$$B_{ij} \stackrel{df}{=} \begin{pmatrix} B_{ij1} \\ \vdots \\ B_{ijq} \end{pmatrix}_{\substack{\text{for } 1 \leq k \leq q \\ \text{such that} \\ Y_{ijk} \neq \phi}} , \quad V_{ij} = \begin{pmatrix} V_{ij1} \\ \vdots \\ V_{ijq} \end{pmatrix}_{\substack{\text{for } 1 \leq k \leq q \\ \text{such that} \\ Y_{ijk} \neq \phi}}$$

$$B_{ijk} \stackrel{df}{=} \begin{pmatrix} B_{ijk1} \\ \vdots \\ B_{ijka_{ijk}} \end{pmatrix} ,$$

$$B_{ijk\beta} \stackrel{df}{=} (0, \ldots, 0, 1, 0, \ldots, 0),$$

is always known, where the 1 in the above 1 by b row vector occurs at only the $b_{ijk\beta}$th position; for outcome

$$T_{ijk} = \{b_{ijk;1}, b_{ijk;2}, \ldots, b_{ijk;a_{ijk}}\},$$

$$a_{ijk} \overset{df}{=} \text{card}\,(T_{ijk}),$$

$$1 \leq b_{ijk;1} < b_{ijk;2} < \cdots < b_{ijk;a_{ijk}} \leq b$$

The V_{ijk}'s are unobserved discrete random variables with known distributions, and are statistically independent of H_i. For different (ijk)'s, the V_{ijk}'s are statistically independent

Given a $Q^{(j)}$, B_i is known.

Unless prior information is available, the probability distribution function for any H_i is not really known to the observer. Consequently for purposes of implementing the proposed score in a real-world setting (where the target identifications are not known), it is assumed initially that all of the H_i's are statistically independent with respect to different i's and possess identical uniform prior distributions over C. On the other hand, for a simulator-dependent measure, the H_i's may be assigned different known prior distributions (including possible dirac ones), if both the attribute properties of each target are known and the index i is identified with the proper target by the simulator.

Thus, we define here

$$
\begin{aligned}
& \mathrm{pr}(H_i = Y^{\sim} \mid Q^{(j)}) \\
&= \mathrm{pr}(H_i = Y^{\sim}) \\
&= \begin{cases} 1/a, & \text{if } Y^{\sim} \in C \\ 0, & \text{if } Y^{\sim} \notin C \end{cases}
\end{aligned}
\tag{2.4}
$$

((8)) **Total Observed Data and Additional Notation**

By assumptions and from subsection ((6)) it follows that

$$\begin{aligned} &\mathrm{pr}(Z^{(j)} \mid Y^{(j)}, Q^{(j)}) \qquad (2.5)\\ &= \mathrm{pr}(Z^{(j)} \mid Q^{(j)})\\ &= \mathrm{pr}(Z_+^{(j)} \mid Q^{(j)}) \cdot \mathrm{pr}(Z_0^{(j)} \mid Q^{(j)}), \end{aligned}$$

noting again that since $Q^{(j)}$ is given, so are the outcomes of the $T_{i\alpha k}$'s.

By suitable rearrangements, the total observed data vector may be broken up into geolocation and non-geolocation atrribute components, and further into geolocation observed target data, geolocation observed false alarm data and non-geolocation attribute data:

$$Z^{(j)} = \begin{pmatrix} Z^{(j)} \\ \hline Y^{(j)} \end{pmatrix} = \begin{pmatrix} Z_+^{(j)} \\ \hline Z_0^{(j)} \\ \hline Y^{(j)} \end{pmatrix} \qquad (2.6)$$

Using the Z notation, the following interpretations hold:

$Z_i^{(j)}$ = track set i (i.e., all data corresponding to $Q_i^{(j)}$) up to t_j,

Z_{ijk} = report (ijk)

= all data for track set i at t_j, from sensor system k,

Z_{ij} = $(Z_{ijk})_{1 \leq k \leq q}$,

Z_j = all data at t_j,

$Z^{(j)}$ = all data up to t_j, etc.

Finally, notation will be introduced which emphasizes at each new sampling time just prior to forming a new updated partitioning (which will incorporate and be based upon the new data), the observer's lack of knowledge:

$$Z_j = (\mathfrak{z}_{\gamma,j,k})_{\substack{\gamma = 1,2,\ldots,m_{jk} \\ k = 1,\ldots,q}} \tag{2.7}$$

up to suitable rearrangments of Z_j, where $m_{jk} \geqq 0$ is the total number of data reports observed by sensor k at t_j - <u>always</u> <u>known</u> - where, recalling from subsection ((3)), a single report consists of either (geolocation) data $\mathfrak{z}'_{\gamma jk}$ and/or non-geolocation attribute data $\mathfrak{z}''_{\gamma jk}$, for sensor k at t_j.

$\mathfrak{z}_{\gamma jk}$ is of the form

$$\mathfrak{z}_{\gamma jk} = Z_{ijk} \quad (r_{jk} \text{ by } 1) \tag{2.8a}$$

for some corresponding unknown (to the observer) $i = i(\gamma,j,k)$, $1 \leqq i \leqq M^{(j)}$, ($M^{(j)}$ also unknown), or it is of the form

$$\mathfrak{z}_{\gamma jk} = Z_{0jk\omega} = Z_{0jk\omega} \quad (r_{jk} \text{ by } 1) \tag{2.8b}$$

where $\omega = \omega(\gamma,j,k)$, $1 \leqq \omega \leqq f_{jk}$, the total number of false alarms observed by sensor system k at t_j. (Recall that $Z_{0jk} = \{Z_{0jk\omega} \mid 1 \leqq \omega \leqq f_{jk}\}$. At most one report $\mathfrak{z}_{jk}$ corresponds to each target for sensor k at t_j; the remaining reports are false alarms.)

Note than for any $Q^{(j)}$, m_{jk} does not depend on $Q^{(j)}$, but does satisfy the relation

$$m_{jk} = q_{(jk)} + f_{jk} + \sum_{\substack{i=1,2,\ldots,M^{(j)} \\ \text{such that} \\ Y_{ijk} \neq \phi, \text{ but } Z_{ijk} = \phi}} 1 \tag{2.9}$$

where $q_{(jk)}$ is the number of target reports seen at t_j by sensor k (see eq. (4.37), f_{jk} is the number of false alarm reports seen at t_j by sensor k (see subsection ((3))), and the last term represents the number of reports seen at t_j by sensor k, which contain only non-geolocation information (and no geolocation information). Unless $Q^{(j)}$ is given, each of the terms on the right hand side of the above equation are unknown to the observer.

3. BASIC DECISION AND INFORMATION THEORETIC PROPERTIES OF THE PROPOSED SCORE

One basic score J for a given correlator-tracker is defined here to be

$$J(\overset{\circ}{Q}^{(j)}, \overset{\circ}{Z}^{(j)}) \stackrel{df}{=\!=} -2 \log pr(Q^{(j)} = \overset{\circ}{Q}^{(j)} | Z^{(j)} = \overset{\circ}{Z}^{(j)}) \qquad (3.1)$$

where $\overset{\circ}{Q}^{(j)}$ is that outcome of partitioning of data corresponding to the correlator-tracker in question and $\overset{\circ}{Z}^{(j)}$ is the observed data outcome.

The corresponding simulator-observer dependent measure is defined to be

$$I(\overset{\circ}{Q}^{(j)}, \overset{\circ\circ}{Q}^{(j)}) \stackrel{df}{=\!=} E_{Z^{(j)}} (J(\overset{\circ}{Q}^{(j)}(Z^{(j)}), Z^{(j)}) | Q^{(j)} = \overset{\circ\circ}{Q}^{(j)}), \qquad (3.2)$$

where $\overset{\circ\circ}{Q}^{(j)}(Z^{(j)})$ is an outcome of the true (but unknown to the observer) partitioning of the data.

Thus $I(\ )$ is a cross entropy measure. (See, e.g., Ref. [22].)

Consider the statistical decision theory game with parameter and decision space being the set of all possible outcomes $\overset{\circ}{Q}^{(j)}$ of $Q^{(j)}$. Observed data $\overset{\circ}{Z}^{(j)}$ has distributions of all relevant random quantities determined from the previous assumptions. For this game, the loss function L is of the zero-one type: for decision outcome $\overset{\circ}{Q}^{(j)}$ and true parameter value $\overset{\circ\circ}{Q}^{(j)}$ of $Q^{(j)}$,

$$L(\overset{\circ}{Q}^{(j)}, \overset{\circ\circ}{Q}^{(j)}) = \begin{cases} 1 \text{ iff } \overset{\circ}{Q}^{(j)} \neq \overset{\circ\circ}{Q}^{(j)} \\ 0 \text{ iff } \overset{\circ}{Q}^{(j)} = \overset{\circ\circ}{Q}^{(j)} \end{cases} \qquad (3.3)$$

Then the Bayes decision function for this game is identifiable with the posterior maximum likelihood estimator $\hat{Q}^{(j)} = \hat{Q}^{(j)}(\overset{\circ}{Z}^{(j)})$, i.e.,

$$\max_{\left(\substack{\text{over all}\\ \overset{\circ}{Q}^{(j)}}\right)} \mathrm{pr}(Q^{(j)} = \overset{\circ}{Q}^{(j)} \mid Z^{(j)} = \overset{\circ}{Z}^{(j)}) \tag{3.4}$$

occurs for $\overset{\circ}{Q}^{(j)} = \hat{Q}^{(j)}$, uniquely.

(See Ref. [23], Chapter 11 and Ref. [24] for background and elaboration of results.)

Thus it immediately follows that

$$\min_{\left(\substack{\text{over all}\\ \overset{\circ}{Q}^{(j)}}\right)} J(\overset{\circ}{Q}^{(j)}, \overset{\circ}{Z}^{(j)})$$

occurs for $\overset{\circ}{Q}^{(j)} = \hat{Q}^{(j)}$, uniquely. (3.5)

If the above statistical decision game is modified to choosing between two hypotheses (the results are readily extended to more than two)

$$H_{(1)}: Q^{(j)} = \overset{\circ}{Q}^{(j)} \text{ vs. } H_{(2)}: Q^{(j)} = \overset{\circ\circ}{Q}^{(j)}$$

with loss function and prior distribution specified by (assuming $\overset{\circ}{Q}^{(j)} \neq \overset{\circ\circ}{Q}^{(j)}$)

$$\begin{aligned} L(\overset{\circ}{Q}^{(j)}, \overset{\circ\circ}{Q}^{(j)}) &\overset{\mathrm{df}}{=} L^{(12)} \\ L(\overset{\circ}{Q}^{(j)}, \overset{\circ}{Q}^{(j)}) &\overset{\mathrm{df}}{=} L^{(11)} \\ L(\overset{\circ\circ}{Q}^{(j)}, \overset{\circ}{Q}^{(j)}) &\overset{\mathrm{df}}{=} L^{(21)} \\ L(\overset{\circ\circ}{Q}^{(j)}, \overset{\circ\circ}{Q}^{(j)}) &\overset{\mathrm{df}}{=} L^{(22)}; \\ L^{(12)}, L^{(21)} &> L^{(11)}, L^{(22)} \end{aligned} \tag{3.6}$$

$$\mathrm{pr}(Q^{(j)} = \overset{\circ}{Q}{}^{(j)}) \stackrel{\mathrm{df}}{=} p_1$$

$$\mathrm{pr}(Q^{(j)} = \overset{\circ\circ}{Q}{}^{(j)}) \stackrel{\mathrm{df}}{=} p_2 = 1 - p_1 \quad , \tag{3.7}$$

then the Bayes decision function for this game is

$$\text{Decide } H_{(1)} \text{ iff } \frac{\mathrm{pr}(Z^{(j)} = \overset{\circ}{Z}{}^{(j)} | Q^{(j)} = \overset{\circ}{Q}{}^{(j)})}{\mathrm{pr}(Z^{(j)} = \overset{\circ}{Z}{}^{(j)} | Q^{(j)} = \overset{\circ\circ}{Q}{}^{(j)})} > T_0'$$

$$\text{Decide } H_{(2)} \text{ iff } \frac{\mathrm{pr}(Z^{(j)} = \overset{\circ}{Z}{}^{(j)} | Q^{(j)} = \overset{\circ}{Q}{}^{(j)})}{\mathrm{pr}(Z^{(j)} = \overset{\circ}{Z}{}^{(j)} | Q^{(j)} = \overset{\circ}{Q}{}^{(j)})} \leqq T_0' \quad , \tag{3.8}$$

where threshold

$$T_0' \stackrel{\mathrm{df}}{=} \frac{p_2}{p_1} \cdot \frac{L^{(21)} - L^{(22)}}{L^{(12)} - L^{(11)}} \tag{3.9}$$

Equivalently, taking logs (see the definition of J' in eq. (3.14))

Decide $H_{(1)}$ iff

$$J'(\overset{\circ}{Q}{}^{(j)}, \overset{\circ}{Z}{}^{(j)}) - J'(\overset{\circ\circ}{Q}{}^{(j)}, \overset{\circ}{Z}{}^{(j)}) < -2 \log T_0'$$

Decide $H_{(2)}$ iff

$$J'(\overset{\circ}{Q}{}^{(j)}, \overset{\circ}{Z}{}^{(j)}) - J'(\overset{\circ\circ}{Q}{}^{(j)}, \overset{\circ}{Z}{}^{(j)}) \geqq -2 \log T_0' \quad . \tag{3.10}$$

(For $p_1 = p_2 = \frac{1}{2}$ and $L_{12} = L_{21}$ and $L_{11} = L_{22}$, $T_0 = 1$, $-2 \log T_0 = 0$.)

Note the equivalent decisions

Decide $H_{(1)}$ iff

$$J(\overset{\circ}{Q}^{(j)},\overset{\circ}{Z}^{(j)}) < J(\overset{\circ\circ}{Q}^{(j)},\overset{\circ}{Z}^{(j)}) - 2 \log T_0$$

Decide $H_{(2)}$ iff

$$J(\overset{\circ}{Q}^{(j)},\overset{\circ}{Z}^{(j)}) > J(\overset{\circ\circ}{Q}^{(j)},\overset{\circ}{Z}^{(j)}) - 2 \log T_0 \tag{3.11}$$

where

$$T_0 \overset{df}{=} \frac{L^{(21)} - L^{(22)}}{L^{(12)} - L^{(11)}} \tag{3.12}$$

For a discussion of the distribution of the J''s and J's, and hence the type I and type II decision errors, as well as the probabilities of correct decisions, see subsection 5.

Suppose now the prior distribution of $Q^{(j)}$ is uniform over the set of, say, γ_j possible outcomes.

Then

$$J(\overset{\circ}{Q}^{(j)},\overset{\circ}{Z}^{(j)}) = J'(\overset{\circ}{Q}^{(j)},\overset{\circ}{Z}^{(j)}) + D_j(\overset{\circ}{Z}^{(j)}) \tag{3.13}$$

where

$$J'(\overset{\circ}{Q}^{(j)},\overset{\circ}{Z}^{(j)}) \overset{df}{=} -2 \log \mathrm{pr}(Z^{(j)} = \overset{\circ}{Z}^{(j)} | Q^{(j)} = \overset{\circ}{Q}^{(j)}) \tag{3.14}$$

and

$$D_j(\overset{\circ}{Z}^{(j)}) \overset{df}{=} 2 \log (\gamma_j \cdot \mathrm{pr}(Z^{(j)} = \overset{\circ}{Z}^{(j)})) \tag{3.15}$$

Thus J and J' differ by a function of $\overset{\circ}{Z}^{(j)}$ and not $\overset{\circ}{Q}^{(j)}$.

Hence, in this case, defining $\hat{\overset{\circ}{Q}}^{(j)} = \hat{\overset{\circ}{Q}}^{(j)}(\overset{\circ}{Z}^{(j)})$ as the conditional maximum likelihood estimator of $Q^{(j)}$, i.e.,

$$\max_{\left(\substack{\text{over all} \\ \overset{\circ}{Q}^{(j)}}\right)} \mathrm{pr}(Z^{(j)} = \overset{\circ}{Z}^{(j)} | Q^{(j)} = \overset{\circ}{Q}^{(j)}) \tag{3.16}$$

occurs for $\overset{\circ}{Q}^{(j)} = \hat{\overset{\circ}{Q}}^{(j)}$, uniquely, then equations (3.5) and (3.13) imply

$$\min_{\left(\substack{\text{over all}\\ \overset{\circ}{Q}^{(j)}}\right)} J(\overset{\circ}{Q}^{(j)}, \overset{\circ}{Z}^{(j)}) \quad \text{and} \quad \min_{\left(\substack{\text{over all}\\ \overset{\circ}{Q}^{(j)}}\right)} J'(\overset{\circ}{Q}^{(j)}, \overset{\circ}{Z}^{(j)}) \tag{3.17}$$

occur for $\overset{\circ}{Q}^{(j)} = \hat{Q}^{(j)} = \hat{\hat{Q}}^{(j)}$, uniquely.

In addition, eq. (3.13) implies that

$$\mathcal{J}(\overset{\circ}{Q}^{(j)}, \overset{\circ\circ}{Q}^{(j)}) = \mathcal{J}'(\overset{\circ}{Q}^{(j)}, \overset{\circ\circ}{Q}^{(j)}) + D'_j(\overset{\circ\circ}{Q}^{(j)}) \tag{3.18}$$

where

$$\mathcal{J}'(\overset{\circ}{Q}^{(j)}, \overset{\circ\circ}{Q}^{(j)}) \overset{\text{df}}{=} E_{Z^{(j)}}(J'(\overset{\circ}{Q}^{(j)}(Z^{(j)}), Z^{(j)} \mid Q^{(j)} = \overset{\circ\circ}{Q}^{(j)}) \tag{3.19}$$

and

$$D'_j(\overset{\circ\circ}{Q}^{(j)}) \overset{\text{df}}{=} E_{Z^{(j)}}(D_j(Z^{(j)}) \mid Q^{(j)} = \overset{\circ\circ}{Q}^{(j)}), \tag{3.20}$$

the latter being a function of $\overset{\circ\circ}{Q}^{(j)}$ only.

Then

$$\min_{\left(\substack{\text{over all}\\ \overset{\circ}{Q}^{(j)}}\right)} \mathcal{J}(\overset{\circ}{Q}^{(j)}, \overset{\circ\circ}{Q}^{(j)}) \quad \text{and} \quad \min_{\left(\substack{\text{over all}\\ \overset{\circ}{Q}^{(j)}}\right)} \mathcal{J}'(\overset{\circ}{Q}^{(j)}, \overset{\circ\circ}{Q}^{(j)}) \tag{3.21}$$

occur uniquely for the same function $\overset{\circ}{Q}^{(j)} = \hat{Q}^{(j)} = \hat{\hat{Q}}^{(j)}$, by using equations (3.17) and (3.18). Also, the Fundamental Inequality of Information Theory (see, e.g., Ref. [22], Chapters 2 and 3) implies that if the $\overset{\circ}{Q}^{(j)}$'s are restricted to constants (not non-trivial functions of $Z^{(j)}$), in the two equivalent minimizations in equation (3.21), the corresponding mimina occur for the common value $\overset{\circ}{Q}^{(j)} = \overset{\circ\circ}{Q}^{(j)}$. These minima are easily seen (via eq. (3.4), e.g.) to have larger values in general than those for the unrestricted case in (3.21).

In summary, the above equations imply that if correlator-tracker has partitioning output $\overset{\bullet}{Q}^{(j)}$ operating on data $\overset{\bullet}{Z}^{(j)}$ and the true but unknown partitioning of the data is actually outcome $\overset{\bullet\bullet}{Q}^{(j)}$ then the lower the score $J(\overset{\bullet}{Q}^{(j)}, \overset{\bullet}{Z}^{(j)})$ - noting, if the uniform prior distribution assumption for $Q^{(j)}$ is made, J can be replaced by J' - the better $\overset{\bullet}{Q}^{(j)}$ approximates $\overset{\bullet\bullet}{Q}^{(j)}$ in both decision and information theory contexts.

4. COMPUTABLE STRUCTURE OF THE SCORE

Throughout this subsection it will be assumed that the prior distribution of $Q^{(j)}$ is uniform so that, without loss of generality, we consider only the score J' and its expectation (with respect to data $Z^{(j)}$) I'. For ease of notation here, outcomes of random quantities - such as $\mathring{Q}^{(j)}$ with respect to $Q^{(j)}$ (partitioning) $\mathring{Z}^{(j)}$ with respect to $Z^{(j)}$ (observed data), etc. - unless ambiguities arise, will be identified with the corresponding random quantities.

The structure of J' is seen to be, using eq. (2.5) and the calculus of conditional and joint probabilities:

$$\begin{aligned} J'(Q^{(j)}, Z^{(j)}) = &-2 \log \mathrm{pr}(Z_+{}^{(j)} | Q^{(j)}) \\ &-2 \log \mathrm{pr}(Z_0{}^{(j)} | Q^{(j)}) \\ &-2 \log \mathrm{pr}(Y^{(j)} | Q^{(j)}), \end{aligned} \tag{4.1}$$

a decomposition into a sum of three terms: the first involving geolocation target track data, the second pertaining to false alarms, and the last relating to the non-geolocation target attribute data.

In turn, each of the terms in eq. (4.1) may be decomposed by straightforward use of the special notation developed in subsection 2, and eqs. (2.1) - (2.4), with the corresponding assumptions:

$$\begin{aligned} &-2 \log \mathrm{pr}(Z_+{}^{(j)} | Q^{(j)}) \\ &= -2 \log \prod_{\begin{pmatrix} 1 \leq i \leq M^{(j)} \\ \text{such that} \\ Z_i{}^{(j)} \neq \phi \end{pmatrix}} \prod_{\begin{pmatrix} 0 \leq \alpha \leq j \\ \text{such that} \\ Z_{i\alpha} \neq \phi \end{pmatrix}} \mathrm{pr}(Z_{i\alpha} | Z_i{}^{(\alpha-1)}, Q^{(j)}) \\ &= \sum_{\begin{pmatrix} 1 \leq i \leq M^{(j)}, \\ \text{such that} \\ Z_i{}^{(j)} \neq \phi \end{pmatrix}} L_i{}^{(j)} \quad , \end{aligned} \tag{4.2}$$

is the measure of overall goodness of fit of the geolocation data for partitioning $Q^{(j)}$ to all the track sets (determined by $Q^{(j)}$).

$(Z_{i\alpha} | Z_i^{(\alpha-1)}, Q^{(j)})$ represents the random variable of goodness of fit of the geolocation data for partitioning $Q^{(j)}$ to track set i at t_α (based on previous data), and

$$L_{i\alpha} = L_{i\alpha}(Z_+^{(\alpha)}, Q^{(j)})$$
$$\underline{\underline{df}} -2 \log \mathrm{pr}(Z_{i\alpha} | Z_i^{(\alpha-1)}, Q^{(j)}) \qquad (4.3)$$

is a measure of goodness of fit of geolocation data at t_α to the i^{th} track set, with respect to $Q^{(j)}$.

$L_i^{(j)}$ is the i^{th} track set goodness of fit of the geolocation data for $Q^{(j)}$ (corresponding to target i, the index i being unknown to the observer), and is given by

$$L_i^{(j)} = L_i^{(j)}(Z_+^{(j)}, Q^{(j)})$$
$$\underline{\underline{df}} -2 \log \mathrm{pr}(Z_i^{(j)} | Q^{(j)})$$
$$= \sum_{\begin{pmatrix} 0 \leqq \alpha \leqq j \\ \text{such that} \\ Z_{i\alpha} \neq \phi \end{pmatrix}} L_{i\alpha} \quad . \qquad (4.4)$$

For $\alpha = 0$, $(Z_{i\alpha} | Z_i^{(\alpha-1)}, Q^{(j)})$ may be interpreted as either $(Z_{i0} | Q^{(j)})$ or using information $Z_i^{(-1)}$ as $(Z_{i0} | Z_i^{(-1)}, Q^{(j)})$.

If $Z_{i\alpha} \neq \phi$ and $Z_i^{(\alpha-1)} \neq \phi$, then since $X_i^{(\alpha)}$, $V_i^{(\alpha)}$ and hence $Z_i^{(\alpha)}$ are all normally distributed (see subsections 2((5)), 2((6)), $(Z_{i\alpha}|Z_i^{(\alpha-1)}, Q^{(j)})$ is distributed normally

$$N_{s_{i\alpha}}(\mu_{i\alpha}^{(j)}, \Sigma_{i\alpha}^{(j)}) , \qquad (4.5)$$

where

$$\begin{aligned}\mu_{i\alpha}^{(j)} &= \mu_{i\alpha}^{(j)}(Z_i^{(\alpha)}, Q^{(j)}) \\ &\overset{df}{=} E(Z_{i\alpha}|Z_i^{(\alpha-1)}, Q^{(j)}) \\ &= B_{\underset{i}{\alpha}} \cdot \hat{X}_{i;\alpha-1,\alpha}^{(j)} ,\end{aligned} \qquad (4.6)$$

$$\begin{aligned}\Sigma_{i\alpha}^{(j)} &= \Sigma_{i\alpha}^{(j)}(Q^{(j)}) \\ &= Cov(Z_{i\alpha}|Z_i^{(\alpha-1)}, Q^{(j)}) \\ &= B_{\underset{i}{\alpha}} \cdot \Lambda_{i,\alpha-1,\alpha}^{(j)} \cdot B_{\underset{i}{\alpha}}^T + R_{\underset{i}{\alpha}} .\end{aligned} \qquad (4.7)$$

$\mu_{i\alpha}^{(j)}$ can be considered to be the optimal estimator (since, e.g., it is both the minimal expected normed square and posterior maximum likelihood estimator) of $X_{i,\alpha}$ based on data $Z^{(\alpha-1)}$, for given partitioning $Q^{(j)}$. In addition, it also can be shown that $\Sigma_{i,\alpha}^{(j)}$ is the covariance matrix of error: $\Sigma_{i\alpha}^{(j)} = Cov(\mu_{i\alpha}^{(j)} - Z_{i\alpha})$. (See Ref. [25] for these and other related results. For convenience, the dependence on j, for $\alpha \leq j$ is not indicated for the $\hat{X}$, Λ and related terms below.)

$$\begin{aligned}\hat{X}_{i;\alpha-1,\alpha} &= \hat{X}_{i;\alpha-1,\alpha}(Z_i^{(\alpha-1)}, Q^{(j)}) \\ &\overset{df}{=} E(X_{i,\alpha}|Z_i^{(\alpha-1)}, Q^{(j)}) \\ &= \text{optimal estimator of} \\ &\quad X_{i\alpha} \text{ given } Z_i^{(\alpha-1)} \text{ and } Q^{(j)},\end{aligned} \qquad (4.8)$$

$$\Lambda_{i;\alpha-1,\alpha} = \Lambda_{i;\alpha-1,\alpha}(Q^{(\alpha)})$$

$$\stackrel{df}{=} \mathrm{Cov}(X_{i\alpha} | Z_i^{(\alpha-1)}, Q^{(j)})$$

$$= \mathrm{Cov}(\hat{X}_{i;\alpha-1,\alpha} - X_{i,\alpha}) \tag{4.9}$$

= covariance matrix of error between $\hat{X}_{i;\alpha-1,\alpha}$ and $X_{i\alpha}$.

$\hat{X}_{i,\alpha-1,\alpha}$ and $\Lambda_{i;\alpha-1,\alpha}$ may be obtained recursively as outputs of the standard Kalman filter (for linear Gauss-Markov data measurement and target motion models - which is the case here; see, e.g. Ref. [26]). For completeness, these equations are presented below:

Define for all $j \geq \alpha \geq 0$, for given $Q^{(j)}$

$$\hat{X}_{i;\alpha,\alpha} \stackrel{df}{=} E(X_{i,\alpha} | Z_i^{(\alpha)}, Q^{(j)}) \tag{4.10}$$

and

$$\Lambda_{i;\alpha,\alpha} \stackrel{df}{=} \mathrm{Cov}(\hat{X}^{(j)}_{i;\alpha,\alpha} - X_{i,\alpha})$$

$$= \mathrm{Cov}(X_{i,\alpha} \mid Z_i^{(\alpha)}, Q^{(j)}) \tag{4.11}$$

Then in the notation developed here, for given $Q^{(j)}$; $0 \leq \alpha \leq j$

$$\hat{X}_{i;\alpha,\alpha+1} = \Phi_{\alpha+1} \cdot \hat{X}_{i;\alpha,\alpha} \quad , \tag{4.12}$$

$$\Lambda_{i;\alpha,\alpha+1} = \Phi_{\alpha+1} \cdot \Lambda_{i;\alpha,\alpha} \cdot \Phi_{\alpha+1}^T + G_{\alpha+1} \cdot P_{\alpha+1} \cdot G_{\alpha+1}^T, \tag{4.13}$$

$$\hat{X}_{i;\alpha,\alpha} = \hat{X}_{i;\alpha-1,\alpha} + K_{i;\alpha} \cdot \nu_{i;\alpha} \ , \tag{4.14}$$

$$K_{i\alpha} = K_{i,\alpha}(Q^{(j)}) \stackrel{df}{=} \begin{cases} \Lambda_{i;\alpha-1,\alpha} \cdot B_{i\alpha}^T \cdot \Sigma_{i,\alpha}^{-1}, & \text{if } Z_{i,\alpha} \neq \phi \\ 0 & , \text{if } Z_{i,\alpha} = \phi \end{cases} \tag{4.15a}$$

$$= \begin{cases} \Lambda_{i,\alpha,\alpha} \cdot B_{\alpha}^{T}{}_{i} \cdot R_{\alpha}^{-1}{}_{i} , & \text{if } Z_{i,\alpha} \neq \phi \\ 0 , & \text{if } Z_{i,\alpha} = \phi \end{cases} \tag{4.15b}$$

$$\nu_{i;\alpha} = \nu_{i,\alpha} (Z_i^{(\alpha)}, Q^{(j)}) \quad \text{(innovations)}$$

$$\underline{\underline{\mathrm{df}}} \; Z_{i,\alpha} - \mu_{i,\alpha} \quad , \tag{4.16}$$

$$\Lambda_{i;\alpha,\alpha} = \begin{cases} (B_{\alpha}^{T}{}_{i} \cdot R_{\alpha}^{-1}{}_{i} \cdot B_{\alpha}{}_{i} + \Lambda_{i;\alpha-1,\alpha}^{-1})^{-1}, & \text{if } Z_{i\alpha} \neq \phi \\ \Lambda_{i,\alpha-1,\alpha} & , \text{if } Z_{i\alpha} = \phi \end{cases} \tag{4.17a}$$

$$= \begin{cases} (I_m - K_{i,\alpha} \cdot B_{\alpha}{}_{i}) \cdot \Lambda_{i;\alpha-1,\alpha} , & \text{if } Z_{i\alpha} \neq \phi \\ \Lambda_{i;\alpha-1,\alpha} & , \text{if } Z_{i\alpha} = \phi \end{cases} \tag{4.17b}$$

Then combining the results of (4.3), (4.5) and (4.16)

$$L_{i\alpha} = L'_{i\alpha} + L''_{i\alpha} \quad , \tag{4.18}$$

where

$$L'_{i\alpha} = L'_{i\alpha} (Q^{(j)})$$

$$\underline{\underline{\mathrm{df}}} \; s_{i\alpha} \cdot \log 2 + \log \det \Sigma_{i\alpha} \tag{4.19}$$

= non-random goodness of fit geolocation data term for track set i at t_{α}, with respect to $Q^{(j)}$

and (noting $s_{i\alpha} = s_{i\alpha}(Q^{(j)})$

$$L''_{i\alpha} = L''_{i\alpha}(Z_i^{(\alpha)}, Q^{(j)})$$

$$\underline{\underline{df}}\ \nu_{i\alpha}^{T} \cdot \Sigma_{i\alpha}^{-1} \cdot \nu_{i\alpha} \qquad (4.20)$$

= random (quadratic in innovations) goodness of fit geolocation data term for track set i at t_α with respect to $Q^{(j)}$; $0 \leq \alpha \leq j$

It should be noted that a nontrivial trade-off exists between the magnitudes of $L'_{i\alpha}$ and $L''_{i\alpha}$. Thus, for the same data set $Z_+^{(j)}$, if $\mathring{Q}^{(j)}$ is a partition with many relatively small (or one point) track sets and $\mathring{\mathring{Q}}^{(j)}$ consists of fewer but bigger track sets, it is not clear which is larger in general: $-2\log pr(Z_+^{(j)}|\mathring{Q}^{(j)})$ or $-2\log pr(Z_+^{(j)}|\mathring{\mathring{Q}}^{(j)})$. In particular, disregarding the innovations $\nu_{i\alpha}$'s and the sizes of the track sets (proportional to the $s_{i\alpha}$'s) - which also really play a role in these trade-offs for $\mathring{Q}^{(j)}$, the $\Sigma_{i\alpha}$'s will tend to be larger (in an eigenvalue or matrix ordering sense) than those in $\mathring{\mathring{Q}}^{(j)}$, due to the $Cov(X_{i\alpha}|Z_i^{(\alpha-1)}, \mathring{Q}^{(j)})$'s being larger (less data reduction for the same target); thus the $L'_{i\alpha}$'s in general may be larger for $\mathring{Q}^{(j)}$ than for $\mathring{\mathring{Q}}^{(j)}$. Yet this implies also that the $\Sigma_{i\alpha}^{-1}$'s for $\mathring{Q}^{(j)}$ will be smaller and hence the $L''_{i\alpha}$'s for $\mathring{Q}^{(j)}$ will be smaller than for $\mathring{\mathring{Q}}^{(j)}$.

In order to evaluate $L'_{i\alpha}$ and $L''_{i\alpha}$, the determinant and multiplicative matrix inverse of $\Sigma_{i\alpha}$ must be obtained. The evaluations of these quantitites, as well as the implementation of the Kalman filter equations - (4.14) - (4.17), in particular - are made more efficient by considering the relative sizes of $s_{i,\alpha}$ and m, for each α; as a consequence, matrix inversion and determinant operations can be applied to matrices no larger than $\min(s_{i,\alpha}, m)$

in dimension.

Specifically, if $s_{i,\alpha} > m$, then for (4.17), (4.17a) may be used; for (4.15), (4.15b) may be used resulting in a simplification for (4.14). In addition identities for $\Sigma_{i\alpha}^{-1}$ and $\log\det \Sigma_{i\alpha}$ can be used. (See, e.g., Ref. [25], Chapter 1.) Thus, multiplying out the block forms for the matrices involved (assuming $Z_{i\alpha} \neq \phi$)

$$\begin{aligned}\Lambda_{i;\alpha,\alpha} &= (D_{i,\alpha} + \Lambda_{i;\alpha-1,\alpha}^{-1})^{-1} \\ &= \Lambda_{i;\alpha-1,\alpha}\cdot(I_m + D_{i\alpha}\cdot\Lambda_{i;\alpha-1,\alpha})^{-1} \\ &= D_{i,\alpha}^{-1}\cdot\{I_m - (I_m + D_{i\alpha}\cdot\Lambda_{i;\alpha-1,\alpha})^{-1}\}\end{aligned} \tag{4.21}$$

where

$$\begin{aligned}D_{i\alpha} &\stackrel{df}{=} B_{i^\alpha}^{T}\cdot R_{i^\alpha}^{-1}\cdot B_{i^\alpha} \\ &= \sum_{\substack{1\leq\beta\leq q \\ \text{such that} \\ Z_{i\alpha\beta}\neq\phi}} B_{\alpha\beta}^{T}\cdot R_{\alpha\beta}^{-1}\cdot B_{\alpha\beta}\ ,\end{aligned} \tag{4.22}$$

$$\begin{aligned}\hat{x}_{i;\alpha,\alpha} &= \hat{x}_{i;\alpha-1,\alpha} + \\ &\quad \Lambda_{i;\alpha,\alpha}\cdot\sum_{\substack{1\leq\beta\leq q \\ \text{such that} \\ Z_{i,\alpha,\beta}\neq\phi}} (B_{\alpha,\beta}^{T}\cdot R_{\alpha,\beta}^{-1}\cdot \nu_{i;\alpha,\beta})\end{aligned} \tag{4.23}$$

where

$$\nu_{i,\alpha,\beta} \stackrel{df}{=} Z_{i,\alpha,\beta} - \mu_{i,\alpha,\beta}\ , \tag{4.24}$$

(innovations for target i by sensor system β at t_α)

$$\mu_{i,\alpha,\beta} = E(Z_{i,\alpha,\beta}|Z^{(\alpha-1)}, Q^{(j)})$$

$$= B_{\alpha,\beta} \cdot \hat{X}_{i;\alpha-1,\alpha} \quad , \tag{4.25}$$

$$\Sigma^{-1}_{i,\alpha} = R^{-1}_{i^\alpha} \cdot (I_{s_{i,\alpha}} - B_{i^\alpha} \cdot K_{i,\alpha})$$

$$= \left(A_{i;\alpha;\beta',\beta''}\right)_{\left(\begin{array}{l} 1 \leq \beta',\beta'' \leq q \\ \text{such that} \\ Z_{i,\alpha,\beta'} \neq \phi, \\ Z_{i,\alpha,\beta''} \neq \phi \end{array}\right)} \tag{4.26}$$

where

$$A_{i;\alpha;\beta',\beta''} \stackrel{df}{=} \delta_{\beta',\beta''} \cdot R^{-1}_{\alpha\beta'} - R^{-1}_{\alpha\beta'} \cdot B_{\alpha\beta'} \cdot \Lambda_{i;\alpha,\alpha} \cdot B^T_{\alpha\beta''} \cdot R^{-1}_{\alpha\beta''} \tag{4.27}$$

$\delta_{\beta',\beta''}$ is the Kronecker delta function (i.e., $\delta_{\beta',\beta''} = 0$, if $\beta' \neq \beta''$; $\delta_{\beta',\beta''} = 1$, if $\beta' = \beta''$).

Hence, (4.20) becomes

$$L''_{i\alpha} = \sum_{\left(\begin{array}{l} 1 \leq \beta',\beta'' \leq q \\ \text{such that} \\ Z_{i\alpha\beta'} \neq \phi, \\ Z_{i,\alpha\beta''} \neq \phi \end{array}\right)} \nu^T_{i;\alpha,\beta'} \cdot A_{i;\alpha;\beta',\beta''} \cdot \nu_{i;\alpha,\beta''}$$

$$= \sum_{\left(\begin{array}{l} 1 \leq \beta \leq q \\ \text{such that} \\ Z_{i,\alpha,\beta} \neq \phi \end{array}\right)} \nu^T_{i;\alpha,\beta} \cdot R^{-1}_{\alpha\beta} \cdot \nu_{i;\alpha,\beta}$$

$$- 2 \sum_{\left(\begin{array}{l} 1 \leq \beta' < \beta'' \; q, \\ \text{such that} \\ Z_{i\alpha\beta'} \neq \phi \\ Z_{i\alpha\beta''} \neq \phi \end{array}\right)} \nu^T_{i;\alpha,\beta'} \cdot R^{-1}_{\alpha\beta'} \cdot B_{\alpha\beta'} \cdot \Lambda_{i;\alpha,\alpha} \cdot B^T_{\alpha\beta''} \cdot R^{-1}_{\alpha\beta''} \nu_{i,\alpha,\beta''} \tag{4.28}$$

Also,

$$\log\det \Sigma_{i\alpha} = \log\det R_{i\alpha}$$

$$+ \log\det (I_m + D_{i\alpha}\cdot\Lambda_{i;\alpha-1,\alpha}), \tag{4.29}$$

noting that

$$\log\det R_{i\alpha} = \sum_{\substack{1 \leq \beta \leq q \\ \text{such that} \\ Z_{i\alpha\beta} \neq \phi}} \log\det R_{\alpha\beta} \;. \tag{4.30}$$

This yields a direct evaluation for $L'_{i\alpha}$ in (4.19).

On the other hand, if $s_{i\alpha} \leq m$, then for (4.17), (4.17b) is appropriate, for (4.15), (4.15a) is preferable, resulting in (4.14) simplifying somewhat:

$$\Lambda_{i,\alpha,\alpha} = (I_m - \Lambda_{i;\alpha-1,\alpha}\cdot F_{i\alpha})\cdot\Lambda_{i;\alpha-1,\alpha} \;, \tag{4.31}$$

where

$$F_{i\alpha} \overset{\text{df}}{=} \sum_{\substack{1 \leq \beta',\beta'' \leq q, \\ \text{such that} \\ Z_{i\alpha\beta'} \neq \phi \\ Z_{i\alpha\beta''} \neq \phi}} B^T_{\alpha\beta'}\cdot A_{i;\alpha;\beta',\beta''}\cdot B_{\alpha\beta''} \tag{4.32}$$

and

$$\hat{X}_{i;\alpha,\alpha} = \hat{X}_{i;\alpha-1,\alpha} + \Lambda_{i;\alpha-1,\alpha}\cdot \sum_{\substack{1 \leq \beta',\beta'' \leq q \\ \text{such that} \\ Z_{i\alpha\beta'} \neq \phi \\ Z_{i\alpha\beta''} \neq \phi}} B^T_{\alpha\beta'}\cdot A_{i;\alpha;\beta',\beta''}\cdot\nu_{i;\alpha,\beta''} \tag{4.33}$$

where the $A_{i;\alpha;\beta',\beta''}$'s ($r_{\alpha\beta'}$ by $r_{\alpha\beta''}$) are determined from the block decomposition of $\Sigma^{-1}_{i\alpha}$:

$$\Sigma_{i\alpha}^{-1} \stackrel{df}{=} (A_{i;\alpha;\beta',\beta''})_{\begin{pmatrix} 1 \le \beta',\beta'' \le q \\ \text{such that} \\ Z_{i\alpha\beta'} \ne \phi, \\ Z_{i\alpha\beta''} \ne \phi \end{pmatrix}} , \tag{4.34}$$

where in block form

$$\Sigma_{i\alpha} = (G_{i;\alpha;\beta',\beta''})_{\begin{pmatrix} 1 \le \beta',\beta'' \le q \\ \text{such that} \\ Z_{i\alpha\beta'} \ne \phi, \\ Z_{i\alpha\beta''} \ne \phi \end{pmatrix}} ; \tag{4.35}$$

with

$$G_{i;\alpha;\beta',\beta''} \stackrel{df}{=} \delta_{\beta',\beta''} \cdot R_{\alpha\beta'} + B_{\alpha\beta'} \cdot \Lambda_{i;\alpha-1,\alpha} \cdot B^T_{\alpha\beta''} . \tag{4.36}$$

The main difficulties in the evaluations of $L'_{i\alpha}$ and $L''_{i\alpha}$ for either relation between $s_{i\alpha}$ and m are as *follows*:

For $s_{i\alpha} > m$, eq. (4.21), e.g., requires matrix inversion (of size m) and eq. (4.29) requires a computation of the log determinant for an m by m matrix (the second term).

For $s_{i\alpha} \le m$, $\Sigma_{i\alpha}^{-1}$ must be obtained, preferably in the specified block form of the $A_{i;\alpha;\beta',\beta''}$'s (see eqs. (4.34) and (4.35)) and log det $\Sigma_{i\alpha}$ must also be evaluated ($\Sigma_{i\alpha}$ is of dimension $s_{i\alpha}$).

In Appendix A, two iterative techniques are presented which can help in resolving these computational difficulties. The first is a procedure for obtaining the inverse in prescribed block form of a positive definite matrix also given in the same form (for a *general number* of blocks). The second, similarly, obtains the log determinant of a positive definite matrix in block form.

Some simplifications for the computations occur when $B_{\alpha,\beta} \equiv B_{\alpha 0}$ and $R_{\alpha\beta} \equiv R_{\alpha 0}$, i.e. are constants for all β such that $Z_{i\alpha\beta} \neq \phi$. (This is not too unreasonable an assumption, since this still allows the sensor systems independent operations, but with the same general measurement characteristics.):

Let

$$q_{i,j} \stackrel{df}{=} \sum_{\substack{1 \leq k \leq q \\ \text{such that} \\ Z_{i,jk} \neq \phi}} 1 \quad . \tag{4.37}$$

Then

$$D_{i\alpha} = q_{i\alpha} \cdot B_{\alpha 0}^T \cdot R_{\alpha 0}^{-1} \cdot B_{\alpha 0} \quad , \tag{4.38}$$

and eq. (4.21) simplifies accordingly;

$$\hat{X}_{i;\alpha,\alpha} = \hat{X}_{i;\alpha-1,\alpha} + q_{i\alpha} \cdot \Lambda_{i;\alpha,\alpha} \cdot B_{\alpha,0}^T \cdot R_{\alpha,0} \cdot \sum_{\substack{1 \leq \beta \leq q \\ \text{such that} \\ Z_{i\alpha\beta} \neq \phi}} \nu_{i;\alpha,\beta} , \tag{4.39}$$

and eqs. (4.24) - (4.28) simplify slightly. Also, (4.38) and

$$\log \det R_{i\alpha} = q_{i\alpha} \cdot \log \det R_{\alpha 0} \tag{4.40}$$

somewhat simplify (4.29). Eqs. (4.31) - (4.33) and (4.36) also simplify slightly.

The special case of one sensor system present (i.e., $q = 1$) shoyld also be noted. In this case, using similar notation as before, either $q_{i\alpha} = 1$ or 0, depending on whether $Z_{i\alpha 1} \neq \phi$ or $Z_{i\alpha 1} = \phi$, respectively. Since $Z_{i\alpha 1} = \phi$ means no geolocation measurement is made of target i at t_α by the only sensor system present,

$L_{i\alpha}$ is not calculated here and in effect $L_{i\alpha} = L'_{i\alpha} = L''_{i\alpha} = 0$. Consider then $z_{i\alpha 1} \neq \phi$. Thus $s_{i\alpha} = r_{i\alpha}$.

For $r_{i\alpha} > m$:

$$D_{i\alpha} = B_{\alpha 1}^{T} R_{\alpha 1}^{-1} B_{\alpha 1} \quad , \tag{4.41}$$

with (4.21) remaining formally the same;

$$\hat{x}_{i;\alpha,\alpha} = \hat{x}_{i;\alpha-1,\alpha} + \Lambda_{i;\alpha,\alpha} B_{\alpha 1}^{T} R_{\alpha 1} \nu_{i;\alpha,1} \quad , \tag{4.42}$$

$$\nu_{i;\alpha,1} = z_{i\alpha 1} - B_{\alpha 1} \cdot \hat{x}_{i;\alpha-1,} \quad , \tag{4.43}$$

$$\begin{aligned} \Sigma_{i\alpha}^{-1} &= A_{i;\alpha;1,1} \\ &= R_{\alpha 1}^{-1} - R_{\alpha 1}^{-1} \cdot B_{\alpha 1} \cdot \Lambda_{i;\alpha,\alpha} \cdot B_{\alpha 1}^{T} \cdot R_{\alpha 1}^{-1} \end{aligned} \tag{4.44}$$

$$L''_{i\alpha} = \nu_{i;\alpha,1}^{T} \cdot A_{i;\alpha;1,1} \cdot \Lambda_{i;\alpha,1} \quad , \tag{4.45}$$

$$L'_{i\alpha} = r_{i\alpha} \log 2\pi + \log \det \Sigma_{i\alpha}, \tag{4.46}$$

$$\log \det \Sigma_{i\alpha} = \log \det R_{\alpha 1} + \log \det (I_m + D_{i\alpha} \cdot \Lambda_{i;\alpha-1,\alpha}) \tag{4.47}$$

For $r_{i\alpha} \leq m$:

$$F_{i\alpha} = B_{\alpha 1}^{T} \cdot A_{i;\alpha;1,1} \cdot B_{\alpha 1} \tag{4.48}$$

with (4.31) remaining formally the same;

$$\hat{x}_{i;\alpha,\alpha} = \hat{x}_{i;\alpha-1,\alpha} + \Lambda_{i;\alpha-1,\alpha} \cdot B_{\alpha 1}^{T} \cdot A_{i;\alpha;1,1} \cdot \nu_{i;\alpha,1} \tag{4.49}$$

equation (4.43) remains the same;

$$\Sigma_{i\alpha}^{-1} = A_{i;\alpha;1,1}$$

$$= (R_{\alpha 1} + B_{\alpha 1}\, \Lambda_{i;\alpha-1,\alpha} B_{\alpha 1}^{T})^{-1} \;; \qquad (4.50)$$

$$\log \det \Sigma_{i\alpha} = \log \det (R_{\alpha 1} + B_{\alpha 1}\, \Lambda_{i;\alpha-1,\alpha}\, B_{\alpha 1}^{T}) \;, \qquad (4.51)$$

with the equations for $L''_{i,\alpha}$ and $L'_{i,\alpha}$ being formally the same as (4.45) and (4.46), respectively.

The next term considered in eq. (4.1), $-2 \log \mathrm{pr}(Z_0^{(j)} \mid Q^{(j)})$, is the goodness of fit of the false alarm data to $Q^{(j)}$ (or vice-versa), $0 \leq \alpha \leq j$

Note first the relations (a consequence of the assumptions in subsections 2((3)) and 2((6)):

$$\mathrm{pr}(Z_0^{(j)} | Q^{(j)}) = \prod_{0 \leq \alpha \leq j} \left(\prod_{\substack{1 \leq k \leq q \\ \text{such that} \\ Z_{0\alpha k} \neq \phi}} \right) \mathrm{pr}(Z_{0\alpha k} | Q^{(j)}) \;, \qquad (4.52)$$

where if $Z_{0\alpha k} \neq \phi$, $f_{\alpha k} > 0$, then

$$\mathrm{pr}(Z_{0\alpha k} | Q^{(j)}) = \prod_{\omega=1}^{f_{\alpha k}} \mathrm{pr}(Z_{0\alpha k\omega} | Q^{(j)})$$

$$= \frac{1}{(2\pi)^{f_{\alpha k} \cdot r_{\alpha k}/2} \cdot (\det M_{\alpha k})^{f_{\alpha k}/2}} \cdot e^{-\frac{1}{2} R_{0\alpha k}} \;; \qquad (4.53)$$

$$R_{0\alpha k} \stackrel{\mathrm{df}}{=\!=} \sum_{\omega=1}^{f_{\alpha k}} (Z_{0\alpha k\omega} - \theta_{\alpha k})^{T} \cdot M_{\alpha k}^{-1} (Z_{0\alpha k\omega} - \theta_{\alpha k})$$

$$= L'_{0\alpha k} + L''_{0\alpha k} \;, \qquad (4.54)$$

$$L'_{0\alpha k} = L'_{0\alpha k}(Z_{0\alpha k}, Q_\alpha)$$

$$\stackrel{df}{=} \sum_{\omega=1}^{f_{\alpha k}} (Z_{0\alpha k} - \bar{Z}_{0\alpha k})^T \cdot M_{\alpha k}^{-1} (Z_{0\alpha k} - \bar{Z}_{0\alpha k})$$

$$= tr(S_{0\alpha k} \cdot M_{\alpha k}^{-1}) , \qquad (4.55)$$

$$S_{0\alpha k} \stackrel{df}{=} \sum_{\omega=1}^{f_{\alpha k}} (Z_{0\alpha k} - \bar{Z}_{0\alpha k}) (Z_{0\alpha k} - \bar{Z}_{0\alpha k})^T , \qquad (4.56)$$

$$\bar{Z}_{0\alpha k} \stackrel{df}{=} (1/f_{\alpha k}) \cdot \sum_{\omega=1}^{f_{\alpha k}} Z_{0\alpha k\omega} \qquad (4.57)$$

$$L''_{0\alpha k} = L''_{0\alpha k} (\bar{Z}_{0\alpha k}, Q^{(j)})$$

$$\stackrel{df}{=} f_{\alpha k} \cdot (Z_{0\alpha k} - \Theta_{\alpha k})^T M_{\alpha k}^{-1} (\bar{Z}_{0\alpha k} - \Theta_{\alpha k}) . \qquad (4.58)$$

(Equation (4.54) follows from the standard procedure of adding and subtracting $\bar{Z}_{0\alpha k}$ within the sum.)

$R_{0\alpha k}$ represents the random (geolocation) fit of the false alarm set for sensor system k at t_α with respect to $Q^{(j)}$, $0 \leq \alpha \leq j$

$L'_{0\alpha k}$ measures the scatter of the false alarm data for sensor system k at t_α with respect to Q_α; $S_{0\alpha k}$ is an unnormalized sample covariance of the false alarm data and $\bar{Z}_{0\alpha k}$ is the sample mean, for sensor k at t_α.

$L''_{0\alpha k}$ measures the bias between the observed and predicted false alarm data means, for sensor k at t_α (with respect to Q_α).

Then, if we define, finally,

$$L'''_{0\alpha k} = L'''_{0\alpha k}(Q^{(j)}) \tag{4.59}$$

$$\stackrel{\text{df}}{=} f_{\alpha k} \cdot (r_{\alpha k} \cdot \log 2\pi + \log \det M_{\alpha k}),$$

a measure of the non-random goodness of fit of the false alarm set for sensor system k at t_α with respect to $Q^{(j)}$, then (4.52) - (4.55), (4.58) and (4.59) yield

$$- 2 \log \mathrm{pr}(Z_0{}^{(j)} | Q^{(j)})$$

$$= \sum_{0 \leq \alpha \leq j} \sum_{\substack{1 \leq k \leq q \\ \text{such that} \\ Z_{0\alpha k} \neq \phi}} L_{0\alpha k} , \tag{4.60}$$

where

$$L_{0\alpha k} = L_{0\alpha k}(Z_{0\alpha k}, | Q^{(j)})$$

$$\stackrel{\text{df}}{=} - 2 \log \mathrm{pr}(Z_{0\alpha k} | Q^{(j)})$$

$$= R_{0\alpha k} + L'''_{0\alpha k}$$

$$= L'_{0\alpha k} + L''_{0\alpha k} + L'''_{0\alpha k}. \tag{4.61}$$

$L_{0\alpha k}$ measures the total goodness of fit of the false alarm set for sensor k at t_α with respect to $Q^{(j)}$, $0 \leq \alpha \leq j$.

Computational problems for the false alarm scores appear minimal; the real difficulty lies in the modeling - specifically in the choice of $\Theta_{\alpha k}$'s and $M_{\alpha k}$'s.

The third term considered in eq. (4.1) $-2\log \mathrm{pr}(Y^{(j)}|Q^{(j)})$, is the likelihood (or goodness-of-fit) of the non-geolocation target attribute data with respect to $Q^{(j)}$.

It follows from the assumptions made in subsection 2((7)) that

$$\mathrm{pr}(Y_i^{(j)}|H_i, Q^{(j)})$$

$$= \prod_{0 \leq \alpha \leq j} \prod_{\begin{pmatrix} 1 \leq k \leq q \\ \text{such that} \\ Y_{i\alpha k} \neq \phi \end{pmatrix}} \mathrm{pr}(Y_{i\alpha k}|H_i, Q^{(j)}). \tag{4.62}$$

Then

$$-2\log \mathrm{pr}(Y^{(j)}|Q^{(j)})$$

$$= -2\log \sum_{\begin{pmatrix} \text{over all} \\ \text{outcomes} \\ \text{of } H^{(j)} \end{pmatrix}} \mathrm{pr}(Y^{(j)}|H^{(j)}, Q^{(j)}) \cdot \mathrm{pr}(H^{(j)})$$

$$= \sum_{\begin{pmatrix} 1 \leq i \leq M^{(j)} \\ \text{such that} \\ Y_i^{(j)} \neq \phi \end{pmatrix}} L_i^{(j)}, \tag{4.63}$$

where

$$L_i^{(j)} = L_i^{(j)}(Y_i^{(j)}, Q^{(j)})$$

$$\underline{\underline{\mathrm{df}}} \; -2\log \mathrm{pr}(Y_i^{(j)}|Q^{(j)})$$

$$= -2\log\Big(\sum_{\begin{pmatrix} \text{over all} \\ \text{outcomes of} \\ H_i \in C \end{pmatrix}} \mathrm{pr}(Y_i^{(j)}|H_i, Q^{(j)}) \cdot \mathrm{pr}(H_i)\Big) \tag{4.64}$$

In addition, note that

$$\mathrm{pr}\left(\begin{pmatrix} Z_+^{(j)} \\ Y^{(j)} \end{pmatrix} \middle| Q^{(j)}\right) = \prod_{\begin{pmatrix} 1 \leq i \leq M^{(j)} \\ \text{such that} \\ Z_i^{(j)} \neq \phi \end{pmatrix}} \mathrm{pr}(Z_i^{(j)}|Q^{(j)}), \tag{4.65}$$

$$\mathrm{pr}(Z_i^{(j)} \mid Q^{(j)}) = \prod_{\substack{0 \leqq \alpha \leqq j \\ \text{such that} \\ Z_{i\alpha} \neq \phi}} \mathrm{pr}(Z_{i\alpha} \mid Z_i^{(\alpha-1)}, Q^{(j)}), \tag{4.66}$$

The goodness of fit of overall data density at time t_α with respect to $Q^{(j)}$ is analogous in form to that of the geolocation only data given in (4.2) et passim, and can be decomposed into a product of two factors, one representing geolocations and the other non-geolocation attributes:

$$\begin{aligned} &\mathrm{pr}(Z_{i\alpha} \mid Z_i^{(\alpha-1)}, Q^{(j)}) \\ &= \mathrm{pr}(Z_{i\alpha} \mid Z_i^{(\alpha-1)}, Q^{(j)}) \cdot \mathrm{pr}(Y_{i\alpha} \mid Y_i^{(\alpha-1)}, Q^{(j)}) \end{aligned} \tag{4.67}$$

(provided $Z_{i\alpha} \neq \phi$ and $Y_{i\alpha} \neq \phi$ for some $0 \leqq \alpha \leqq j$).

Also

$$\begin{aligned} L_i^{(j)} &= -2 \log \prod_{\substack{0 \leqq \alpha \leqq j \\ \text{such that} \\ Y_i \neq \phi}} \mathrm{pr}(Y_{i\alpha} \mid Y_i^{(\alpha-1)}, Q^{(j)}) \\ &= \sum_{\substack{0 \leqq \alpha \leqq j \\ \text{such that} \\ Y_i \neq \phi}} -2 \log \mathrm{pr}(Y_{i\alpha} \mid Y_i^{(\alpha-1)}, Q^{(j)}) \\ &= \sum_{\substack{0 \leqq \alpha \leqq j \\ \text{such that} \\ Y_i \neq \phi}} L_{i\alpha}, \end{aligned} \tag{4.68}$$

where

$$\begin{aligned} L_{i\alpha} &= L_{i\alpha}(Y_i^{(\alpha)}, Q^{(j)}) \\ &\stackrel{\text{df}}{=} -2 \log \mathrm{pr}(Y_{i\alpha} \mid Y_i^{(\alpha-1)}, Q^{(j)}), \end{aligned} \tag{4.69}$$

noting for $0 \leqq \alpha \leqq j$

$$\mathrm{pr}(Y_{i\alpha}|Y_i^{(\alpha-1)}, Q^{(j)}) = \sum_{\substack{\text{over all}\\ \text{outcomes}\\ \text{of } H_i \varepsilon C}} \mathrm{pr}(Y_{i\alpha}|H_i, Y_i^{(\alpha-1)}, Q^{(j)}) \cdot \mathrm{pr}(H_i) \qquad (4.70)$$

If a and b are relatively small integers, by using the (known) distribution functions of the V_{ijk}'s or, equivalently, the $p(Y', Y''; Q^{(j)})$'s, and eq. (2.4) (see subsection 2((7)), then equation (4.63) can be evaluated. However, for relatively large a and b this may require - because of the discrete nature of the distributions involved - a large number of operations. Consequently the following <u>approximation</u> is proposed for the modeling of the random variables Y_{ijk} and H_i, for either relatively large a and b, and/or when the non-geolocation attribute set C is perhaps better modeled as a contiguous subset of b-dimensional Euclidean space:

In equation (2.3b), assume each V_{ijk} (for $Y_{ijk} \neq \phi$) is normally distributed as $N_{a_{ijk}}(0, R_{ijk})$, where R_{ijk} is the submatrix of R, corresponding to outcome T_{ijk} with respect to $\{1,2,.., b\}$, R being a fixed (positive definite) matrix of dimension b, to be determined.

Define, analogous to $R_{\substack{\alpha\\ i}}$,

$$R_{i^\alpha} \stackrel{\mathrm{df}}{=} \begin{pmatrix} R_{i\alpha 1} & & 0 \\ & \ddots & \\ 0 & & R_{i\alpha q} \end{pmatrix} \quad \begin{array}{l}\text{For only}\\ 1 \leq k \leq q\\ \text{such that}\\ Y_{i\alpha k} \neq \phi \end{array}.$$

Replace equation (2.4) by:

H_i is normally distributed as $N_b(E(H_i), Cov(H_i))$, with $E(H_i)$ and $Cov(H_i)$ to be determined in relation to attribute set C. The following trivial identity is used here:

$$H_{i,(\alpha)} \equiv H_i, \quad \text{for all times } t_\alpha . \tag{4.71}$$

From now on for given time t_j and $Q^{(j)}$, the function dependency of quantities on $Q^{(j)}$ will be often omitted, unless ambiguity results or emphasis is desired.

Then to evaluate $L_{i\alpha}$, a Kalman filter can be applied to (2.3b) and (4.71) yielding, analogous to the computations for $L_{i\alpha}$ (see eqs. (4.5) - (4.20), for $0 \leq \alpha \leq j$, for given $Q^{(j)}$:

$$\hat{H}_{i;\alpha,\alpha+1} = \hat{H}_{i;\alpha,\alpha} \tag{4.72}$$

$$\tilde{\Lambda}_{i;\alpha,\alpha+1} = \tilde{\Lambda}_{i;\alpha,\alpha} \tag{4.73}$$

$$\hat{H}_{i;\alpha,\alpha} = \hat{H}_{i;\alpha-1,\alpha} + \mathcal{K}_{i;\alpha} \cdot \tilde{\nu}_{i;\alpha} \tag{4.74}$$

$$\mathcal{K}_{i;\alpha} \stackrel{\text{df}}{=} \begin{cases} \tilde{\Lambda}_{i;\alpha-1,\alpha} \cdot \mathcal{B}_{i\alpha}^T \cdot \tilde{\Sigma}_{i,\alpha}^{-1} & , \text{ if } Y_{i\alpha} \neq \phi \\ 0 & , \text{ if } Y_{i\alpha} = \phi \end{cases} \tag{4.75a}$$

$$= \begin{cases} \tilde{\Lambda}_{i;\alpha,\alpha} \cdot \mathcal{B}_{i\alpha}^T \cdot \mathcal{R}_{i\alpha}^{-1} & , \text{ if } Y_{i\alpha} \neq \phi \\ 0 & , \text{ if } Y_{i\alpha} = \phi \end{cases} \tag{4.75b}$$

$$\tilde{\Lambda}_{i;\alpha,\alpha} = \begin{cases} (\mathcal{B}_{i\alpha}^T \cdot \mathcal{R}_{i\alpha}^{-1} \cdot \mathcal{B}_{i\alpha} + \tilde{\Lambda}_{i;\alpha-1,\alpha}^{-1})^{-1} & , \text{ if } Y_{i\alpha} \neq \phi \\ \tilde{\Lambda}_{i;\alpha-1,\alpha} & , \text{ if } Y_{i\alpha} = \phi \end{cases} \tag{4.76a}$$

$$= \begin{cases} (I_b - K_{i,\alpha} \cdot \underset{i}{B}_{\alpha}) \cdot \tilde{\Lambda}_{i;\alpha-1,\alpha} , & \text{if } Y_{i\alpha} \neq \phi \\ \tilde{\Lambda}_{i;\alpha-1,\alpha} , & \text{if } Y_{i\alpha} = \phi \end{cases} \quad (4.76b)$$

$$\tilde{\nu}_{i;\alpha} \stackrel{df}{=} Y_{i\alpha} - \underset{i}{B}_{\alpha} \cdot \hat{H}_{i;\alpha-1,\alpha} \quad (4.77)$$

$$\tilde{\Sigma}_{i,\alpha} \stackrel{df}{=} \underset{i}{B}_{\alpha} \cdot \tilde{\Lambda}_{i;\alpha-1,\alpha} \cdot \underset{i}{B}_{\alpha}^{T} + \underset{i}{R}_{\alpha} . \quad (4.78)$$

Then

$$L_{i\alpha} = L'_{i\alpha} + L''_{i\alpha} , \quad (4.79)$$

where

$$L'_{i\alpha} \stackrel{df}{=} s_{i\alpha} \cdot \log 2\pi + \log \det \tilde{\Sigma}_{i\alpha} , \quad (4.80)$$

$$s_i \stackrel{df}{=} \sum_{\substack{1 \leq k \leq q \\ \text{such that} \\ Y_{i\,k} \neq \phi}} a_{i\alpha k} , \quad (4.81)$$

$$L''_{i\alpha} \stackrel{df}{=} \tilde{\nu}_{i\alpha}^{T} \cdot \tilde{\Sigma}_{i\alpha}^{-1} \cdot \tilde{\nu}_{i\alpha} . \quad (4.82)$$

Analagous to the cases $s_{i\alpha} > m$ vs. $s_{i\alpha} \leq m$, to evaluate the above equations, we must consider separately the cases $s_{i\alpha} > b$ and $s_{i\alpha} \leq b$:

For $s_{i\alpha} > b$, use eqs. (4.76a) and (4.75b), obtaining

$$\tilde{\Lambda}_{i;\alpha,\alpha} = (D_{i\alpha} + \tilde{\Lambda}_{i;\alpha-1,\alpha}^{-1})^{-1}$$

$$= \tilde{\Lambda}_{i;\alpha-1,\alpha} \cdot (I_b + D_{i\alpha} \cdot \tilde{\Lambda}_{i;\alpha-1,\alpha}) , \quad (4.83)$$

etc.,

$$\mathcal{D}_{i\alpha} \overset{\text{df}}{=} \mathcal{B}_{i^\alpha}^{T} \cdot \mathcal{R}_{i^\alpha}^{-1} \mathcal{B}_{i^\alpha}$$

$$= \sum_{\begin{pmatrix} 1 \leq \beta \leq q \\ \text{such that} \\ Y_{i\alpha\beta} \neq \phi \end{pmatrix}} \mathcal{B}_{i\alpha\beta}^{T} \mathcal{R}_{i\alpha\beta}^{-1} \mathcal{B}_{i\alpha\beta} \quad , \tag{4.84}$$

$$\hat{H}_{i;\alpha,\alpha} = \hat{H}_{i;\alpha-1,\alpha} \quad +$$

$$\tilde{\Lambda}_{i;\alpha,\alpha} \cdot \sum_{\begin{pmatrix} 1 \leq \beta \leq q \\ \text{such that} \\ Y_{i\alpha\beta} \neq \phi \end{pmatrix}} (\mathcal{B}_{i\alpha\beta}^{T} \cdot \mathcal{R}_{i\alpha\beta}^{-1} \cdot \tilde{\nu}_{i;\alpha,\beta}) \quad , \tag{4.85}$$

$$\tilde{\nu}_{i;\alpha\beta} \overset{\text{df}}{=} Y_{i\alpha\beta} - \mathcal{B}_{i\alpha\beta} \cdot \hat{H}_{i;\alpha-1,\alpha} ,$$

$$\tilde{\Sigma}_{i\alpha}^{-1} = \mathcal{R}_{i^\alpha}^{-1} \cdot (I_{\delta_{i\alpha}} - \mathcal{B}_{i^\alpha} \cdot K_{i,\alpha}) ,$$

$$= (A_{i;\alpha;\beta';\beta''})_{\begin{pmatrix} 1 \leq \beta', \beta'' \leq q \\ \text{such that} \\ Y_{i\alpha\beta'} \neq \phi, \\ Y_{i\alpha\beta''} \neq \phi \end{pmatrix}}$$

$$A_{i;\alpha;\beta',\beta''} \overset{\text{df}}{=} \delta_{\beta',\beta''} \cdot \mathcal{R}_{i\alpha\beta'}^{-1} - \mathcal{R}_{i\alpha\beta'}^{-1} \cdot \mathcal{B}_{i\alpha\beta'} \cdot \tilde{\Lambda}_{i;\alpha\alpha} \cdot \mathcal{B}_{i\alpha\beta''}^{T} \cdot \mathcal{R}_{i\alpha\beta''}^{-1} \tag{4.86}$$

$$L_{i\alpha}'' = \sum_{\begin{pmatrix} 1 \leq \beta', \beta'' \leq q \\ \text{such that} \\ Y_{i\alpha\beta'} \neq \phi \\ Y_{i\alpha\beta''} \neq \phi \end{pmatrix}} \tilde{\nu}_{i;\alpha\beta'}^{T} A_{i;\alpha,\beta',\beta''} \tilde{\nu}_{i;\alpha\beta''} \quad , \tag{4.87}$$

etc.

$$\log \det \tilde{\Sigma}_{i\alpha} = \log \det \mathcal{R}_{i^{\alpha}}$$

$$+ \log \det \left(I_b + \mathcal{D}_{i\alpha} \cdot \tilde{\Lambda}_{i,\alpha \cdots ,\alpha}\right), \tag{4.88}$$

$$\log \det \mathcal{R}_{i^{\alpha}} = \sum_{\substack{1 \leq \beta \leq q \\ \text{such that} \\ Y_{i\alpha\beta} \neq \phi}} \log \det \mathcal{R}_{i\alpha\beta} \quad . \tag{4.89}$$

Suppose now $\mathcal{R}$ is in diagonal form:

$$\mathcal{R} \stackrel{\text{df}}{=} \begin{pmatrix} r_i & & 0 \\ & \ddots & \\ 0 & & r_b \end{pmatrix} \tag{4.90}$$

Then

$$\mathcal{R}_{ijk} = \begin{pmatrix} r_{b_{ijk;1}} & & 0 \\ & \ddots & \\ 0 & & r_{b_{ijk;a_{ijk}}} \end{pmatrix} \tag{4,91}$$

recalling

$$T_{ijk} = \{b_{ijk;1}, \ldots, b_{ijk;a_{ijk}}\} \subseteq \{1, \ldots, b\} .$$

Also,

$$\mathcal{B}^T_{i\alpha\beta} \mathcal{R}^{-1}_{i\alpha\beta} \mathcal{B}_{i\alpha\beta} = \begin{pmatrix} d_{i\alpha\beta,1} & & 0 \\ & \ddots & \\ 0 & & d_{i\alpha\beta,b} \end{pmatrix} , \tag{4,92}$$

where

$$d_{i\alpha\beta\gamma} \stackrel{\text{df}}{=} \begin{cases} 0 & , \text{ if } \gamma \notin T_{i\alpha\beta} \\ 1/r_\gamma & , \text{ if } \gamma \in T_{i\alpha\beta} , \end{cases} \tag{4,93}$$

for $1 \leq \gamma \leq b$.

Hence,

$$\mathcal{D}_{i\alpha} = \begin{pmatrix} d_{i\alpha,1} & & 0 \\ & \ddots & \\ 0 & & d_{i\alpha,b} \end{pmatrix}, \tag{4.94}$$

$$d_{i\alpha\gamma} = \sum_{\substack{1 \leq \beta \leq q \\ \text{such that} \\ Y_{i\alpha\beta} \neq \phi}} d_{i\alpha\beta\gamma} = (1/r_\gamma) \cdot \sum_{\substack{1 \leq \beta \leq q, \\ \text{such that } Y_{i\alpha\beta} \neq \phi \\ \text{and there is an } \eta, \\ 1 \leq \eta \leq b, \text{ such that} \\ \gamma = b_{i\alpha\beta\eta} \in T_{i\alpha\beta}}} 1 \tag{4.95}$$

for $1 \leq \gamma \leq b$.

$$\mathcal{B}^T_{i\alpha\beta} \mathcal{R}^{-1}_{i\alpha\beta} \cdot \tilde{\nu}_{i;\alpha,\beta} = \begin{pmatrix} g_{i\alpha\beta:1} \\ \vdots \\ g_{i\alpha\beta;b} \end{pmatrix}, \tag{4.96}$$

where $$g_{i\alpha\beta\gamma} \overset{\text{df}}{=} \begin{cases} 0 & \text{, if } \gamma \notin T_{i\alpha\beta} \\ (1/r_\gamma) \cdot \tilde{\nu}_{i;\alpha;\beta;\eta} & \text{, if } \gamma = b_{i\alpha\beta\eta} \in T_{i\alpha\beta} \end{cases} \tag{4.97}$$

where $$\tilde{\nu}_{i\alpha\beta} = \begin{pmatrix} \tilde{\nu}_{i;\alpha\beta;1} \\ \vdots \\ \tilde{\nu}_{i;\alpha,\beta;a_{i\alpha\beta}} \end{pmatrix} = \begin{pmatrix} Y_{i\alpha\beta;1} - \hat{H}_{i;\alpha-1,\alpha;b_{i\alpha\beta 1}} \\ \vdots \\ Y_{i\alpha\beta;a_{i\alpha\beta}} - \hat{H}_{i;\alpha-1,\alpha;b_{i\alpha\beta a_{i\alpha\beta}}} \end{pmatrix} \tag{4.98}$$

for

$$\hat{H}_{i,\alpha-1,\alpha} \overset{\text{df}}{=} \begin{pmatrix} \hat{H}_{i,\alpha-1,\alpha;1} \\ \vdots \\ \hat{H}_{i,\alpha-1,\alpha;b} \end{pmatrix}, \quad H_i \overset{\text{df}}{=} \begin{pmatrix} H_{i1} \\ \vdots \\ H_{ib} \end{pmatrix}, \tag{4.99}$$

and

$$Y_{i\alpha\beta} \stackrel{df}{=} \begin{pmatrix} Y_{i\alpha\beta;1} \\ \vdots \\ Y_{i\alpha\beta;a_{i\alpha\beta}} \end{pmatrix} \qquad (4.100)$$

Also, letting in scalar form

$$\tilde{\tilde{\Lambda}}_{i;\alpha,\alpha} = (\tilde{\Lambda}_{i;\alpha,\alpha;\gamma',\gamma''})_{1 \leq \gamma',\gamma'' \leq b} \quad , \qquad (4.101)$$

then

$$B_{i\alpha\beta'} \cdot \tilde{\tilde{\Lambda}}_{i;\alpha,\alpha} \cdot B^{T}_{i\alpha\beta''} = (\tilde{\tilde{\Lambda}}_{i;\alpha,\alpha;\gamma',\gamma''})_{\substack{\gamma' \varepsilon T_{i\alpha\beta'} \\ \gamma'' \varepsilon T_{i\alpha\beta''}}} \quad . \qquad (4.102)$$

Note that

$$\log \det R_{i\alpha\beta} = \sum_{n=1}^{a_{i\alpha\beta}} \log r_{b_{i\alpha\beta;n}}$$

$$= \sum_{\gamma \varepsilon T_{i\alpha\beta}} \log r_{\gamma} \quad . \qquad (4.103)$$

For $s_{i\alpha} \leq b$,

$$\tilde{\Lambda}_{i;\alpha,\alpha} = (I_b - \tilde{\Lambda}_{i;\alpha-1,\alpha} \cdot F_{i\alpha}) \cdot \tilde{\Lambda}_{i;\alpha-1,\alpha} \qquad (4.104)$$

where

$$F_{i\alpha} \stackrel{df}{=} \sum_{\begin{pmatrix} 1 \leq \beta',\beta'' \leq q, \\ \text{such that} \\ Y_{i\alpha\beta'} \neq \phi, \\ Y_{i\alpha\beta''} \neq \phi \end{pmatrix}} B^{T}_{i\alpha\beta'} A_{i\alpha\beta',\beta''} B_{i\alpha\beta''} \qquad (4.105)$$

$$\hat{H}_{i;\alpha,\alpha} = \hat{H}_{i;\alpha-1,\alpha} + \tilde{\Lambda}_{i;\alpha-1,\alpha} \cdot \sum_{\substack{1 \leqq \beta',\beta'' \leqq q \\ \text{such that} \\ Y_{i\alpha\beta'} \neq \phi, \\ Y_{i\alpha\beta''} \neq \phi}} B^T_{i\alpha\beta'} \cdot A_{i;\alpha;\beta',\beta''} \cdot \tilde{\nu}_{i;\alpha,\beta''} \tag{4.106}$$

$$\tilde{\Sigma}_{i\alpha} = (G_{i;\alpha;\beta',\beta''})_{\substack{1 \leqq \beta',\beta'' \leqq q, \\ \text{such that} \\ Y_{i\alpha\beta'} \neq \phi, \\ Y_{i\alpha\beta''} \neq \phi}} , \tag{4.107}$$

$$G_{i;\alpha;\beta',\beta''} \stackrel{\text{df}}{=} \delta_{\beta',\beta''} \cdot R_{i\alpha\beta'} + B_{i\alpha\beta'} \cdot \tilde{\Lambda}_{i;\alpha-1,\alpha} \cdot B^T_{i\alpha\beta''} \tag{4.108}$$

Further simplifications can be carried out, since R is in diagonal form.

For example, if

$$A_{i;\alpha;\beta',\beta''} = (a_{i;\alpha;\beta',\beta'';\eta',\eta''})_{\substack{1 \leqq \eta' \leqq a_{i\alpha\beta'} \\ 1 \leqq \eta'' \leqq a_{i\alpha\beta''}}} \tag{4.109}$$

then

$$B^T_{i\alpha\beta'} \cdot A_{i\alpha\beta'\beta''} \cdot B_{i\alpha\beta''} = (h_{i,\alpha,\beta',\beta'';\gamma',\gamma''})_{1 \leqq \gamma',\gamma'' \leqq b} \tag{4.110}$$

where

$$h_{i\alpha\beta',\beta'',\gamma',\gamma''} \stackrel{\text{df}}{=} \begin{cases} 0, \text{ if } \gamma' \notin T_{i\alpha\beta'} \text{ or } \gamma'' \notin T_{i\alpha\beta''} \\ a_{i;\alpha;\beta',\beta'',\eta',\eta''} \quad , \text{ if} \\ \quad \gamma' = b_{i\alpha\beta'\eta'} \varepsilon\, T_{i\alpha\beta'} \quad \& \\ \quad \gamma'' = b_{i\alpha\beta''\eta''} \varepsilon\, T_{i\alpha\beta''} \quad . \end{cases} \tag{4.111}$$

Similarly,

$$B^T_{i\alpha\beta''}\cdot A_{i;\alpha;\beta',\beta''}\cdot\tilde{\nu}_{i;\alpha\beta''}=\begin{pmatrix}P_{i\alpha\beta'\beta''1}\\ \vdots\\ P_{i\alpha\beta'\beta''b}\end{pmatrix}\qquad , \tag{4.112}$$

where

$$P_{i\alpha\beta'\beta''\gamma}=\begin{cases}0\ , \text{ if } \gamma\notin T_{i\alpha\beta'}\\ \displaystyle\sum_{\eta''=1}^{a_{i\alpha\beta''}} a_{i,\alpha,\beta';\beta'',\eta',\eta''}\cdot\tilde{\nu}_{i,\alpha,\beta'',\eta''}\ ,\text{ if}\\ \qquad\qquad\qquad \gamma=b_{\alpha\beta'\eta'}\ \varepsilon\ T_{i\alpha\beta'}\ .\end{cases} \tag{4.113}$$

Alternatively, $L_i^{(j)}$ can be computed directly using eq. (4.71) and the fact that $(Y_i^{(j)}|Q^{(j)})$ is normally distributed with

$$E(Y_i^{(j)}|Q^{(j)}) = B_i^{(j)}\cdot E(H_i) \tag{4.114}$$

and

$$\mathrm{Cov}(Y_i^{(j)}|Q^{(j)}) = B_i^{(j)}\cdot\mathrm{Cov}(H_i)\cdot B_i^{(j)T}+R_i^{(j)}, \tag{4.115}$$

where

$$B_i^{(j)}\ \underline{\underline{df}}\begin{pmatrix}B_{i0}\\ B_{i1}\\ \vdots\\ B_{ij}\end{pmatrix}\quad ;\quad R_i^{(j)}\ \underline{\underline{df}}\begin{pmatrix}R_{i0} & & & 0\\ & R_{i1} & & \\ & & \ddots & \\ 0 & & & R_{ij}\end{pmatrix} \tag{4.116}$$

Let

$$\delta_i^{(j)}\ \underline{\underline{df}}\ \sum_{\alpha=0}^{j}\delta_{i\alpha} \tag{4.117}$$

and assume $s_i^{(j)} \geqq b$. Then

$$\mathcal{L}_i^{(j)} = \mathcal{L}_i^{(j)\prime} + \mathcal{L}_i^{(j)\sim}, \tag{4.118}$$

$$\mathcal{L}_i^{(j)\prime} \underset{=}{\underline{\mathrm{df}}} s_i^{(j)} \cdot \log 2\pi + \log \det \mathrm{Cov}(Y_i^{(j)} | Q^{(j)}) \tag{4.119}$$

$$\mathcal{L}_i^{(j)\sim} \underset{=}{\underline{\mathrm{df}}} (Y_i^{(j)} - E(Y_i^{(j)} | Q^{(j)}))^T \cdot \mathrm{Cov}(Y_i^{(j)} | Q^{(j)})^{-1} \cdot (Y_i^{(j)} - E(Y_i^{(j)} | Q^{(j)})). \tag{4.120}$$

Analogous to eq. (4.29),

$$\log \det \mathrm{Cov}(Y_i^{(j)} | Q^{(j)}) = \log \det \mathcal{R}_i^{(j)} + \log \det (I_b + \mathcal{D}_i^{(j)} \cdot \mathrm{Cov}(H_i)), \tag{4.121}$$

$$\begin{aligned} \mathcal{D}_i^{(j)} \underset{=}{\underline{\mathrm{df}}}\ & \mathcal{B}_i^{(j)T} \cdot \mathcal{R}_i^{(j)-1} \cdot \mathcal{B}_i^{(j)} \\ = & \sum_{\alpha=0}^{j} \mathcal{D}_{i\alpha}, \end{aligned} \tag{4.122}$$

$$\begin{aligned} \mathrm{Cov}(Y_i^{(j)} | Q^{(j)})^{-1} &= (\mathcal{B}_i^{(j)} \mathrm{Cov}(H_i) \mathcal{B}_i^{(j)T} + \mathcal{R}_i^{(j)})^{-1} \\ &= \mathcal{R}_i^{(j)-1} - \mathcal{R}_i^{(j)-1} \cdot \mathcal{B}_i^{(j)} \cdot (\mathrm{Cov}(H_i)^{-1} + \mathcal{D}_i^{(j)})^{-1} \mathcal{B}_i^{(j)T} \cdot \mathcal{R}_i^{(j)-1}, \end{aligned} \tag{4.123}$$

where

$$\mathcal{R}_i^{(j)} \underline{\mathrm{df}} \begin{pmatrix} \mathcal{R}_{i0} & & 0 \\ & \ddots & \\ 0 & & \mathcal{R}_{ij} \end{pmatrix}, \tag{4.124}$$

$$\mathcal{B}_i^{(j)^T} \cdot \mathcal{R}_i^{(j)^{-1}} \cdot (Y_i^{(j)} - E(Y_i^{(j)} | Q^{(j)})) = \sum_{\alpha=0}^{j} \sum_{\substack{1 \leq k \leq q \\ \text{such that} \\ Y_{i\alpha k} \neq \phi}} \mathcal{B}_{i\alpha k}^T \cdot \mathcal{R}_{i\alpha k}^{-1} \cdot (Y_{i\alpha k} - E(Y_{i\alpha k} | Q^{(j)})) \tag{4.125}$$

$$E(Y_{i\alpha} | Q^{(j)}) = \mathcal{B}_{i\alpha} \cdot E(H_i) \; ,$$

and (4.126)

$$E(Y_{i\alpha k} | Q^{(j)}) = \mathcal{B}_{i\alpha k} E(H_i) \; ; \; 1 \leq k \leq q \quad .$$

$$(Y_i^{(j)} - E(Y_i^{(j)} | Q^{(j)}))^T \cdot \mathcal{R}_i^{(j)^{-1}} \cdot (Y_i^{(j)} - E(Y_i^{(j)} | Q^{(j)}))$$

$$= \sum_{\alpha=0}^{j} \sum_{\substack{1 \leq k \leq q \\ \text{such that} \\ Y_{i\alpha k} \neq \phi}} (Y_{i\alpha k} - E(Y_{i\alpha k} | Q^{(j)}))^T \cdot \mathcal{R}_{i\alpha k}^{-1} \cdot (Y_{i\alpha k} - E(Y_{i\alpha k} | Q^{(j)})) \quad . \tag{4.127}$$

Then, $L_i^{(j)\sim}$ can be evaluated by using (4.122) - (4.127) in (4.120).

If, furthermore, $\mathcal{R}$ and $Cov(H_i)$ are diagonal, additional simplifications can be achieved by use of (4.90) - (4.95), replacing $\tilde{\nu}_{i;\alpha,\beta}$ by $Y_{i\alpha\beta} - E(Y_{i\alpha\beta} | Q^{(j)})$ in eqs. (4.96) - (4.100).

An approximation for $L_i^{(j)\sim}$, the computations of which do not depend on $E(H_i)$, which is more accurate for larger (in the positive definite ordering sense) $Cov(H_i)$, is given in Appendix B. Thus, using the approximation no knowledge by the observer is needed of $E(H_i)$ and $Cov(H_i)$, except that for all i, $Cov(H_i)$ exceeds a certain fixed large bound.

When $Cov(H_i)$ is diagonal,

$$\mathrm{Cov}(H_i) = \begin{pmatrix} \mathrm{var}(H_{i1}) & & 0 \\ & \ddots & \\ 0 & & \mathrm{var}(H_{ib}) \end{pmatrix}$$

5. SOME DISTRIBUTIONAL PROPERTIES OF THE SCORE

As in the last subsection, it is assumed that the prior distribution of $Q^{(j)}$ is uniform so that J' and I' need only be considered.

Then from eqs. (4.1), (4.2), (4.4), (4.18), (4.19), (4.20), (4.60), (4.61), (4.59). (4.54) and (4.62) - (4.64) and the assumptions made in subsection 2 $(J'(Q^{(j)},Z^{(j)})|Q^{(j)})$, as a random quantity functionally dependent on $Z^{(j)}$, is distributed as the sum of a constant, a chi-square random variable and a statistically independent discrete valued random variable. Specifically,

$$J'(Q^{(j)},Z^{(j)}) = J_1'(Q^{(j)}) + J_2'(Q^{(j)}, Z^{(j)}) + J_3'(Q^{(j)},Y^{(j)}) \tag{5.1}$$

$$J_1'(Q^{(j)}) \stackrel{\text{df}}{=} \sum_{\substack{1 \leq i \leq M^{(j)} \\ \text{such that} \\ Z_i^{(j)} \neq \phi}} \; \sum_{\substack{0 \leq \alpha \leq j \\ Z_{i\alpha} \neq \phi}} L'_{i\alpha} + \sum_{0 \leq \alpha \leq j} \; \sum_{\substack{1 \leq k \leq q \\ \text{such that} \\ Z_{0\alpha k} \neq \phi}} L'''_{0\alpha k} \tag{5.2}$$

is the constantterm;

$$J_2'(Q^{(j)}), Z^{(j)}) \stackrel{\text{df}}{=} \sum_{\substack{1 \leq i \leq M^{(j)} \\ \text{such that} \\ Z_i^{(j)} \neq \phi}} \; \sum_{\substack{0 \leq \alpha \leq j \\ \text{such that} \\ Z_{i\alpha} \neq \phi}} L''_{i\alpha} + \sum_{0 \leq \alpha \leq j} \; \sum_{\substack{1 \leq k \leq q \\ \text{such that} \\ Z_{0\alpha k} \neq \phi}} R_{0\alpha k} \; , \tag{5.3}$$

(since all $(L_i^{\sim} | Q^{(j)})$'s are distributed as $\chi^2_{s_{i\alpha}}$ statistically independent for all (i,α)'s and statistically independent of all $(R_{0\alpha k} | Q^{(j)})$'s which are distributed as $\chi^2_{f_{\alpha k} \cdot r_{\alpha k}}$) is distributed as $\chi^2_{\xi_j}$, where

$$\xi_j = \xi_j(Q^{(j)}, Z^{(j)})$$

$$\stackrel{df}{=} \sum_{\begin{pmatrix} 1 \leq i \leq M^{(j)} \\ \text{such that} \\ Z_i^{(j)} \neq \phi \end{pmatrix}} \sum_{\begin{pmatrix} 0 \leq \alpha \leq j \\ \text{such that} \\ Z_i \neq \phi \end{pmatrix}} s_{i\alpha}$$

$$+ \sum_{0 \leq \alpha \leq j} \sum_{\begin{pmatrix} 1 \leq k \leq q \\ \text{such that} \\ Z_{0\alpha k} \neq \phi \end{pmatrix}} f_{\alpha k} \cdot r_{\alpha k} \quad . \tag{5,4}$$

$$J_3'(Q^{(j)}, Y^{(j)}) \stackrel{df}{=} \sum_{\begin{pmatrix} 1 \leq i \leq M^{(j)} \\ \text{such that} \\ Y_i^{(j)} \neq \phi \end{pmatrix}} L_i^{(j)} \tag{5.5}$$

has a discrete distribution.

Under the normal approximation made for $(Y^{(j)} | Q^{(j)}$ (see the results following eq. (4.70), especially (4.113) - (4.115)),it follows that the discrete valued random variable $J_3'(Q^{(j)}, Y^{(j)})$ is replaced in the sum comprising $J'(Q^{(j)}, {}^{(j)})$ (eq. (5.1)) by

$$\sum_{\begin{pmatrix} 1 \leq i \leq M^{(j)} \\ \text{such that} \\ Y_i^{(j)} \neq \phi \end{pmatrix}} (L_i^{(j)\prime} + L_i^{(j)\prime\prime}) \qquad \text{(see eq. (4.113)),}$$

which causes an adjusted computation for J'. $J'(Q^{(j)}, Z^{(j)} | Q^{(j)})$ is now distributed as the sum of a constant and a chi-square random variable. Specifically,

$$J'(Q^{(j)}, Z^{(j)}) = J_1''(Q^{(j)}) + J_2''(Q^{(j)}, Z^{(j)}), \tag{5.6}$$

where

$$J_1''(Q^{(j)}) \stackrel{df}{=} J_1'(Q^{(j)}) + \sum_{\substack{1 \leq i \leq M^{(j)} \\ \text{such that} \\ Y_i^{(j)} \neq \phi}} L_i^{(j)'} \tag{5.7}$$

is the constant term ;

$$J_2''(Q^{(j)}, Z^{(j)}) \stackrel{df}{=} J_2'(Q^{(j)}, Z^{(j)}) + \sum_{\substack{1 \leq i \leq M^{(j)} \\ \text{such that} \\ Y_i^{(j)} \neq \phi}} L_i^{(j)''} \tag{5.8}$$

is distributed as $\chi^2_{\eta_j}$, where

$$\begin{aligned} \eta_j &= \eta_j(Q^{(j)}, Z^{(j)}) \\ &\stackrel{df}{=} \xi_j + \sum_{\substack{1 \leq i \leq M^{(j)} \\ \text{such that} \\ Y_i^{(j)} \neq \phi}} \delta_i^{(j)} \quad . \end{aligned} \tag{5.9}$$

It should be noted that if $\overset{\circ}{Q}^{(j)} \neq \overset{\infty}{Q}^{(j)}$, the distribution of $(J'(\overset{\circ}{Q}^{(j)}, Z^{(j)}) | Q^{(j)} = \overset{\circ\circ}{Q}^{(j)})$ is in general not obtainable in a simple closed form - even under the normal approximation made for

$(Y^{(j)}|Q^{(j)})$. This is due mainly to the mismatching of covariance matrices of formally assumed data model by the correlator tracker through $\overset{\circ}{Q}^{(j)}$, with respect to the actual data model through $\overset{\circ\circ}{Q}^{(j)}$.

For example, in reality let $q=1$, $j=5$, $Z_0^{(5)} \equiv \phi$, $Y^{(5)} \equiv \phi$, $\Phi_\alpha \equiv \Phi_0$, $P_\alpha \equiv 0$, $B_{i\alpha} \equiv B_{\alpha 1} \equiv B_{01}$, $R_{i\alpha} \equiv R_{\alpha 1} \equiv R_{01}$. (See eqs. (2.1) and (2.2).) Thus no false alarms nor non-geolocation attribute data are truly present and all observations and target motion are homogeneous and stationary.

Suppose also that only (geolocation) data $Z_{1,\alpha,1} \neq \phi$, for $\alpha = 0, 1, .., 5 = j$ is observed (see the concluding part of subsection 2((8)) for use of notation), where for each α, in general $Z_{1\alpha 1} = Z_{i\alpha 1}$, $i = i(\alpha)$ an unknown positive integer.

Suppose now the observer uses a correlator-tracker determined by partitioning $\overset{\circ}{Q}^{(5)}$, consisting of a single component $\overset{\circ}{Q}_1^{(5)}$, $\overset{\circ}{Q}_1^{(5)} = \{(1,\alpha,1) \mid \alpha = 0, 1, ..,5\}$, corresponding to $Z_1^{(\alpha)} = \{Z_{1,\alpha,1} \mid \alpha = 0, 1, .., 5\}$.

Now *if* $Q^{(5)} = \overset{\circ}{Q}^{(5)}$ *really holds*, then without loss of generality, $Z_{1\cdot\cdot} = Z_{1\alpha 1}$, $\alpha = 0, 1, .., 5$, and the observation and target models in eqs. (2.1), (2.2) combine to:

$$\underbrace{\begin{pmatrix} Z_{101} \\ Z_{111} \\ \vdots \\ Z_{151} \end{pmatrix}}_{= Z^{(5)}} = \underbrace{\begin{pmatrix} B_{01} \\ B_{01}\cdot\Phi_0 \\ \vdots \\ B_{01}\cdot\Phi_0^5 \end{pmatrix}}_{\overset{df}{=} \mathcal{B}} X_{101} + \begin{pmatrix} V_{101} \\ V_{111} \\ \vdots \\ V_{151} \end{pmatrix} \qquad (5.10)$$

yielding ($z^{(5)}$ being normally distributed)

$$E(z^{(5)} | \dot{Q}^{(5)}) = \mathfrak{B} \cdot E(X_{101})$$

$$\mathrm{Cov}(z^{(5)} | \dot{Q}^{(5)}) = \mathfrak{B} \cdot \mathrm{Cov}(X_{101}) \cdot \mathfrak{B}^T + R_1^{(5)}.$$

Now, $J'(\mathring{Q}^{(5)}, z^{(5)}) = L_1^{(5)}(\mathring{Q}^{(5)}, z^{(5)})$ has random term (see eq. (4.20)).

$$L_5 = L_5\ (\mathring{Q}^{(5)}, z^{(5)})$$

$$\underline{\underline{\mathrm{df}}} \sum_{\alpha=0}^{5} L''_{1\alpha}(\mathring{Q}^{(5)}, z^{(5)}) \tag{5.11}$$

$$= \sum_{\alpha=0}^{5} \nu_{12}^T \cdot \Sigma_{12} \cdot \nu_{12}$$

$$= (z^{(5)} - E(z^{(5)}))^T (\mathrm{Cov}(z^{(5)})^{-1} \cdot (z^{(5)} - E(z^{(5)}))$$

(conditioning here is on $\mathring{Q}^{(5)}$).

Then (5.10) and (5.11) imply that $(L_5(\mathring{Q}^{(5)}, z^{(5)}) | \mathring{Q}^{(5)})$ is distributed as $\chi^2_{\xi_5}$, where now

$$\xi_5 = \sum_{\alpha}^{5} s_{i\alpha}. \tag{5.12}$$

On the other hand, <u>suppose</u> $Q^{(5)} = \overset{\infty}{Q}{}^{(5)}$ <u>really holds</u>, where now

$$\overset{\infty}{Q}{}^{(5)} \ \underline{\underline{\mathrm{df}}} \ \{\overset{\infty}{Q}{}_1^{(5)}, \overset{\infty}{Q}{}_2^{(5)}\},$$

$$\overset{\infty}{Q}{}_1^{(5)} = \{(1,\alpha,1) | \alpha = 0, 1, 2\},$$

$$\overset{\infty}{Q}{}_2^{(5)} = \{(2,\alpha,1) | \alpha = 3, 4, 5\}.$$

Thus $\overset{\circ\circ}{Q}_1^{(5)}$ corresponds to $\{Z_{1\alpha 1}|\alpha=0,1,2\} = \{Z_{1\alpha 1}|\alpha=0,1,2\}$ and $\overset{\infty}{Q}_2^{(5)}$ corresponds to $\{Z_{2\alpha 1}|\alpha= 3,4,5\} = \{Z_{1\alpha 1}|\alpha=3,4,5\}$.

This yields from eqs. (2.1), (2.2) the combined observation model:

$$\underbrace{\begin{pmatrix} Z_{101} \\ Z_{111} \\ Z_{121} \\ Z_{231} \\ Z_{241} \\ Z_{251} \end{pmatrix}}_{=Z^{(5)}} = \left(\begin{array}{c|c} \mathbb{B}' & 0 \\ \hline 0 & \mathbb{B}' \end{array}\right) \cdot \begin{pmatrix} X_{101} \\ \\ X_{231} \end{pmatrix} + \begin{pmatrix} V_{101} \\ V_{111} \\ V_{121} \\ V_{231} \\ V_{241} \\ V_{251} \end{pmatrix}, \tag{5.13}$$

where $\mathbb{B}' \stackrel{df}{=} \begin{pmatrix} B_{01} \\ B_{01}\cdot\Phi \\ B_{01}\cdot\Phi^2 \end{pmatrix}$.

But, since here

$E(Z^{(5)}|\overset{\infty}{Q}^{(5)}) = \begin{pmatrix} \mathbb{B}'\cdot E(X_{101}) \\ \mathbb{B}'\cdot E(X_{231}) \end{pmatrix}$, and most importantly

$$Cov(Z^{(5)}|\overset{\infty}{Q}^{(5)}) = \left(\begin{array}{c|c} \mathbb{B}'\cdot Cov(X_{101})\cdot\mathbb{B}'^T & 0 \\ \hline 0 & \mathbb{B}'Cov(X_{231})\mathbb{B}'^T \end{array}\right) + R_1^{(5)} \tag{5.14}$$

does not match in shape (nor size) $Cov(Z^{(5)}|\overset{\circ}{Q}^{(5)})$ which is the central factor in the quadratic form in (5.11), it follows that $(L_5(\overset{\circ}{Q}^{(5)},Z^{(5)}\,|\,\overset{\infty}{Q}^{(5)})$ is not even distributed as a noncentral chi-square random variable; hence neither can $(J'(\overset{\circ}{Q}^{(5)},Z^{(5)})\,|\overset{\infty}{Q}^{(5)})$ have a simple distribution.

Note that, by similar reasoning, if a correlator-tracker is used which formally assumes $\overset{\infty}{Q}^{(5)}$ to be true, then $(L_5(\overset{\infty}{Q}^{(5)},Z^{(5)})|\,\overset{\circ}{Q}^{(5)})$

is distributed as $\chi^2_{\xi_5}$ as in (5.12), but, as before, $(L_5(\overset{\infty}{Q}^{(5)}, Z^{(5)}) | \overset{\infty}{Q}^{(5)})$ has a complicated (non-chi-square, non-noncentral chi-square) distribution in general.

Consequently, in general, for $\overset{\circ}{Q}^{(j)} \neq \overset{\infty}{Q}^{(j)}$, since we cannot obtain a simple computable distribution for $J'(\overset{\circ}{Q}^{(j)}, Z^{(j)} | \overset{\infty}{Q}^{(j)})$, similar remarks hold for the difference $J_j \overset{df}{=} J'(\overset{\circ}{Q}^{(j)}, Z^{(j)}) - J'(\overset{\infty}{Q}^{(j)}, Z^{(j)})$, conditioned on $\overset{\circ}{Q}^{(j)}$. Thus, the computation of a threshold T (see eqs. (3.8) - (3.10) based only on a given significance level of discrimination β (obtained by solving $\beta = \Pr$ (Decide $H_{(2)} | H_{(1)}$ true) $= \Pr(J_j > T | \overset{\infty}{Q}^{(j)})$ for T) appears equally infeasible.

In summary, $J'(\overset{\circ}{Q}^{(j)}, \overset{\circ}{Z}^{(j)})$ <u>can</u> be used two ways in a real-world situation:

(1) For determining how well a given correlator-tracker, through its partitioning of data $\overset{\circ}{Q}^{(j)}$ really fits data $\overset{\circ}{Z}^{(j)}$, by evaluating the cumulative distribution function of the random quantity $(J\ (\overset{\circ}{Q}^{(j)}, Z^{(j)}) | \overset{\circ}{Q}^{(j)})$ at the outcome point $J\ (\overset{\circ}{Q}^{(j)}, \overset{\circ}{Z}^{(j)})$. This distribution in general should be computable, since the random quantity here is the statistically independent sum of a computable constant, a chi-square random variable with a computable number of degrees of freedom and a discrete random variable which has computable characteristics (theoretically at least, if parameters a and b are small) (see eqs. (5.2) - (5.5).) (An approximation based on simplifying the computational burden for the discrete random variable term is given in eqs. (5.6) - (5.9).

(2) For comparing two (or more) given correlator-tracker schemes, operating on the same data. In general, up to an adjustment for

different decision losses or gains and presence of prior distributional information on $Q^{(j)}$ (see eqs. (3.6) - (3.10)), that correlator tracker, determined by partitioning $Q^{(j)}$, is chosen among a given set $\{\overset{\circ}{Q}^{(j)}, \tilde{Q}^{(j)}, .., \overset{\approx}{Q}^{(j)}\}$ say, for which $J'(Q^{(j)}, \overset{\circ}{Z}^{(j)})$ is minimal. However, the statistical significance of the differences between the values of $J'(Q^{(j)}, \overset{\circ}{Z}^{(j)})$ for different correlator-trackers $Q^{(j)}$ is apparently difficult to determine.

Note that using the distributional results in eqs. (5.2) - (5.5), for example, $I'(\overset{\circ}{Q}^{(j)}, \overset{\circ}{Q}^{(j)})$ is easily obtained as

$$\mathfrak{I}'(\overset{\circ}{Q}^{(j)}, \overset{\circ}{Q}^{(j)}) = J_1' \; (Q^{(j)}) + \xi_j + E(J_3'(Q^{(j)}, Y^{(j)})). \qquad (5.15)$$

If the normal approximation is made (for $Y^{(j)}|Q^{(j)}$), then eqs. (5.6) - (5.9) imply

$$\mathfrak{I}'(\overset{\circ}{Q}^{(j)}, \overset{\circ}{Q}^{(j)}) = J_1''(Q^{(j)}) + \eta_j \qquad (5.16)$$

$\mathfrak{I}'$ evaluated as in (5.16) can be used also as measure of <u>average</u> performance of $\overset{\circ}{Q}^{(j)}$ (with respect to averaging the data $Z^{(j)}$) and can be used analagously to J' in (2.) for comparing average goodness of fit of several correlator-trackers in question, by choosing that one minimizing the corresponding value of $\mathfrak{I}'$.

SUMMARY

A simple scoring rule for correlator-trackers is developed in this report. Mathematical-logical justifications for the use of this score are demonstrated. Detailed computations necessary to implement the rule are exhibited, along with a determination of its statistical distribution, a form of the chi-square.

In the second part of the study, numerical examples will be given illustrating the suitability of the scoring technique for use in surveillance in a real world setting and leading to sensitivity analysis with respect to the key parameters involved in correlation. Future work will concentrate on both extending the applicability of the score and obtaining further analytic properties.

REFERENCES

1. I. R. Goodman, *Preliminary Study on Error Bounding and Related Aspects of the Correlation Problem*, NRL Memorandum 7930-85: IRG, Naval Research Laboratory, Wash., D. C., March 21, 1980.

2. D. Reid, *Multiple Hypotheses Filter for Tracking Multiple Targets in a Cluttered Environment* Report LMSC-D560254, Sept., 1977. Shorter version appears in *PROC. 17TH IEEE Conf. on Decision and Control*, San Diego, CA, (1202-1211) Jan., 1979.

3. C. L. Morefield, "Decision Directed Multi-Target Tracking", *Proc. 17th Conf. on Decision and Control*, San Diego, CA, (1195-1201) Jan., 1979. Also see the earlier "Applications of 0-1 Integer Programming to Multi-Target Tracking", *IEEE Trans., on Auto. Control*, Vol. AC-22, No. 3, (302-312) June, 1977.

4. H. L. Wiener, W. W. Willman, J. H. Kullback and I. R. Goodman, *Naval Ocean Surveillance Correlation Handbook (CY-1978)*, NRL Report 8340, Naval Research Laboratory, Wash., D.C., Oct. 31, 1979. The Handbook for FY-1979 is in press (Goodman, et al.)

5. J. H. Kullback and M. E. Owens, *Multisensor Correlation for Ocean Surveillance: Problems and Limitations*, NRL Report 7258, Naval Research Laboratory, Wash., D.C., May 22, 1972 (CONFIDENTIAL).

6. R. W. Sittler, "Optimal Data Association Problem in Surveillance Theory", *IEEE Trans. Mil. Elec.*,Vol. MIL-8, (125-139) April 1964.

7. I. R. Goodman, *A General Model for the Contact Correlation Problem*, NRL report in press (Spring, 1980). Shorter version appears in *Proc. 18th IEEE Conf. on Decision and Control*, Ft. Lauderdale, Fla., Dec., 1979.

8. Y. Bar-Shalom, "Tracking Methods in a Multitarget Evnironment", *IEEE Trans. Auto. Control*, Vol. AC-23, No. 4, (618-626) Aug., 1978.

9. W. W. Willman, *Some Performance Results for Recursive Multitarget Correlation-Tracker Algorithms*,NRL report in press, April 1980, Naval Research Laboratory, Wash., D. C.,

10. H. L. Wiener, *Tracking Algorithm Performance Measures*, NRL Memo - internal (7930-169), Naval Research Laboratory, Wash., D. C., June 20, 1975 (Confidential).

11. H. L. Wiener, A. S. Distler, and J. H. Kullback, "Operational and Implementation Problems of Multi-Target Correlator-Trackers", *Proc., 18th IEEE Conf. on Dec. and Control*,Ft. Lauderdale, Fla., Dec., 1979.

12. R. Blinkenberg, Briefing to NRL personnel, Aug. 10, 1979. Slides for SCAT (Sea Automatic Tracker). Aerospace Corp., El Segundo, CA.

13. KETRON, Inc., *Track-Data Handling Model Study Analysis* IR 2366, KETRON, Inc., Arlington, VA, April 1, 1975.

14. M. A. Adler et al.., *Correlation Tracking Performance Study - DCP Input*, Report R-1650, Vols. I, II, PRC Systems Sciences Co., June 15, 1973 (Secret).

15. R. D. Turner and S. Marder, *The Track Establishment Process in a Mixed False Alarm-Rate Environment*, I.D.A., Arlington, VA, Paper P-728, June, 1972.

16. A. L. Quigley, *A Theory For Simple Associations During Target Tracking*, Tech. Report TR-71-11, Admiralty Surface Weapons Establish., Portsmouth, England, Feb., 1971

17. A. L. Quigley, *A Theory for Complex Association*, Tech. Report TR-72-13, Admiralty Surface Weapons Establish., Portsmouth, England, April, 1972.

18. N. Wax, "Signal to Noise Improvement and the Statistics of Track Populations", *J. of Applied Physics*, Vol. 26, No. 5, (586-595) May, 1955.

19. P. W. Stephan, *Software Development Study*, Report No. P73-242, Hughes Aircraft Co., Culver City, CA, June 1973 (Secret). Updated by Personal Communication, Sept., 1979.

20. D. L. Alspach and R. N. Lobbia, "A Score for Correct Data Association in Multi-Target Tracking", *Proc. 18th IEEE Conf. on Dec. and Control*, Ft. Lauderdale, Fla., Dec., 1979.

21. J. J. Stein and S. S. Blackman, "Generalized Correlation of Multi-Target Track Data", *IEEE Trans. Aero. and Elec. Sys.*, Vol. AES-11, No. 6, (1207-1217) Nov., 1975.

22. S. Kullback, *Information Theory and Statistics*, New York: John Wiley & Sons, Inc., 1959 (Later rereleased as a Dover book).

23. D. Blackwell and M. A. Girshick, *Theory of Games and Statistical Decisions*, New York: John Wiley & Sons, Inc., 1954.

24. D. Middleton, *An Introduction to Statistical Communication Theory*, New York: McGraw-Hill Book Co., Inc., 1960. Section 4.

25. C. R. Rao, *Linear Statistical Inference and Its Applications*, 2nd Ed., New York: John Wiley & Sons, Inc., 1973.

26. A. H. Jazwinski, *Stochastic Processes & Filtering Theory*, New York: Academic Press (Math. in Sci. & Eng. Vol. 64), 1970.

Appendix A

Matrix Inverse and Log Determinant of a Positive Definite Matrix in Prescribed Block Form

Let $G = (G_{ij})_{1 \leq i,j \leq n}$ be a given positive definite matrix in block form with all G_{ii} assumed square. Thus all G_{ii}'s are positive definite with $G_{ij}^T = G_{j,i}$, etc.

Let $G^{-1} \stackrel{df}{=} (A_{ij})_{1 \leq i,j \leq n}$, with the A_{ij}'s to be determined.

Then

1. For $1 \leq i \leq j \leq n$

$$A_{ji} = (\delta_{ij} \cdot I - \sum_{k=i+1}^{n} A_{jk} \cdot G_{ki}^{(i)}) \cdot G_{ii}^{(i)\,-1}, \qquad (A.1)$$

assuming the sum is zero for i=n, where for any $1 \leq i,j,\ k \leq n$, $G_{ik}^{(j)} = G_{ki}^{(j)T}$.

For $i = 1$, compute for $j = 1,2,\ \ j \leq k \leq n$,

$$G_{kj}^{(1)} \stackrel{df}{=} G_{kj} \qquad (A.2)$$

For i such that $n \geq i > 1$, compute for $j = 1, i+1,\ \ j \leq k \leq n$. (For $i = n$, let $k = j = i = n$.)

$$G_{kj}^{(i)} \stackrel{df}{=} G_{jk}^{(i-1)} - G_{k,i-1}^{(i-1)} \cdot G_{i-1,i-1}^{(i-1)^{-1}} \cdot G_{i-1,j}^{(i-1)} \qquad (A.3)$$

2. $$\log \det G = \sum_{i=1}^{n} \log \det (G_{ii}^{(i)}) \ . \tag{A.4}$$

For i such that $n \geq i \geq 1$ compute $G_{kj}^{(i)}$ for $i+1 \geq k \geq j \geq i$. (For $i = n$, let $k = j = i = n$.)

The proofs of the above results easily follow by successively applying the matrix and determinantal identities in Ref. [25], Chapter 1, pp. 32, 33.

Appendix B

An Approximation to $\mathcal{L}_i^{(j)\sim}$

Assume rank $(\mathcal{B}_i^{(j)}) = b < \delta^{(j)}$. Define

$$\tilde{\mathcal{L}}_i^{(j)\sim} \stackrel{df}{=} Y_i^{(j)^T} \cdot \mathcal{H}_i^{(j)} \cdot Y_i^{(j)} ,$$

where

$$\mathcal{H}_i^{(j)} \stackrel{df}{=} \mathcal{R}_i^{(j)^{-1}} - \mathcal{R}_i^{(j)^{-1}} \cdot \mathcal{B}_i^{(j)} \cdot \mathcal{D}_i^{(j)^{-1}} \cdot \mathcal{B}_i^{(j)^T} \cdot \mathcal{R}_i^{(j)^{-1}}$$

Let

$$h_i \stackrel{df}{=} \text{mineig}\ (\text{Cov}(H_i))$$

$$h_{ij1} \stackrel{df}{=} \text{mineig}\ \text{Cov}(\mathcal{R}_{ij})$$

$$h_{ij2} \stackrel{df}{=} \text{mineig}\ (\mathcal{D}_i^{(j)}) \quad .$$

Then, for any arbitrary given $\varepsilon > 0$, if

$$h_i > (1/h_{ij2}) \cdot \max(0, (\|Y_i^{(j)} - E(Y_i^{(j)}\|^2/\varepsilon \cdot h_{ij1}) - 1),$$

then (see eq. (4.120))

$$0 \leq \mathcal{L}_i^{(j)\sim} - \tilde{\mathcal{L}}_i^{(j)\sim} < \varepsilon \quad .$$

Proof:

Note that by using matrix identities (Ref. [25], Chapter 1; see also eq. (4.123))

$$Cov^{-1}(Y_i^{(j)}) - \mathcal{H}_i^{(j)} = \mathcal{R}_i^{(j)^{-1}} \cdot \Omega_i^{(j)} \cdot \mathcal{R}_i^{(j)^{-1}}, \tag{B-1}$$

where (see eq. (4.117)

$$\Omega_i^{(j)} = \mathcal{B}_i^{(j)} \cdot (\mathcal{D}_i^{(j)} \cdot Cov(H_i) \cdot \mathcal{D}_i^{(j)} + \mathcal{D}_i^{(j)})^{-1} \cdot \mathcal{B}_i^{(j)^T} . \tag{B-2}$$

Then

$$\mathcal{R}_i^{(j)^{-1}} \cdot \Omega_i^{(j)} \cdot \mathcal{R}_i^{(j)^{-1}}$$

$$= \mathcal{R}_i^{(j)^{-\frac{1}{2}}} \cdot (\mathcal{R}_i^{(j)^{-\frac{1}{2}}} \cdot \Omega_i^{(j)} \cdot \mathcal{R}_i^{(j)^{-\frac{1}{2}}}) \cdot \mathcal{R}_i^{(j)^{-\frac{1}{2}}} . \tag{B.3}$$

Since

$$\mathcal{R}_i^{(j)^{-\frac{1}{2}}} \cdot \Omega_i^{(j)} \cdot \mathcal{R}_i^{(j)^{-\frac{1}{2}}}$$

$$\leqq \text{maxeig } (\mathcal{R}_i^{(j)^{-\frac{1}{2}}} \cdot \Omega_i^{(j)} \cdot \mathcal{R}_i^{(j)^{-\frac{1}{2}}}) \cdot I_{\delta_i^{(j)}} \tag{B-4}$$

(in positive definite matrix ordering sense).

Now

$$\text{maxeig } (\mathcal{R}_i^{(j)^{-\frac{1}{2}}} \cdot \Omega_i^{(j)} \cdot \mathcal{R}_i^{(j)^{-\frac{1}{2}}})$$

$$= \text{maxeig } ((\mathcal{D}_i^{(j)} Cov(H_i)\ \mathcal{D}_i^{(j)} + \mathcal{D}_i^{(j)})^{-1} \cdot \mathcal{D}_i^{(j)})$$

$$= \text{maxeig } (Cov(H_i) \cdot \mathcal{D}_i^{(j)} + I_n)^{-1}$$

$$= 1/(\text{mineig } (Cov(H_i) \cdot \mathcal{D}_i^{(j)}) + 1)$$

$$\leqq 1/(h_i \cdot h_{ij2} + 1) \tag{B-5}$$

Noting that $H_i^{(j)} \cdot B_i^{(j)} = 0$, it follows that

$H_i^{(j)} \cdot E(Y_i^{(j)}) = H_i^{(j)} \cdot B_i^{(j)} \cdot E(H_i) = 0$ and thus combining (B-1) - (B-5)

$$0 \leq (Y_i^{(j)} - E(Y_i^{(j)}))^T \cdot (Cov^{-1}(Y_i^{(j)}) - H_i^{(j)}) \cdot (Y_i^{(j)} - E(Y_i^{(j)})$$

(omitting the conditioning on $Q^{(j)}$)

$$= L_i^{(j)''} - \tilde{L}_i^{(j)''}$$

$$\leq \|Y_i^{(j)} - E(Y_i^{(j)}\|^2 \cdot \text{maxeig}\ (R_i^{(j)^{-\frac{1}{2}}} \cdot (1/(h_i h_{ij2} + 1)) \cdot I \cdot R_i^{(j)^{-\frac{1}{2}}})$$

$$\leq \|Y_i^{(j)} - E(Y_i^{(j)}\|^2 \cdot (1/(h_{ij1}(h_i h_{ij2} + 1))) \tag{B-6}$$

Appendix C

Flow Charts for Computations of the Score J′

The number of sensor systems is q, known and fixed.

At each time t_j, new raw data (possibly vacuous) Z_j is observed before $Q^{(j)}$ is determined. $Z_j = (z_{\gamma,j,k})_{\substack{\gamma = 1,\ldots,m_{jk} \\ k = 1.,,,q}}$.

For each sensor system k, at t_j, $m_{jk} \geqq 0$ is the known number of data reports $z_{\gamma,j,k}$ observed. (When $m_{jk} = 0$, data becomes missing, i.e., $z_{\gamma jk} = \phi$ when $m_{jk} \geqq 1$, only at most one report can correspond to each true target - the remaining are false alarms.) Also, $z_{\gamma jk}$ is decomposed into a possible geolocation (target or false alarm - not known) data component $z'_{\gamma jk}$, which if nonvacuous is r_{jk} by 1, $r_{jk} \geqq 1$ *known, and a possible non-geolocation* target data component, $z''_{\gamma jk}$, which if nonvacuous, rules out the associated geolocation component being a false alarm, and is of dimension $\leqq b$; C is the fixed known non-geolocation target attribute set of b by 1 vectors, each representing a feasible evaluation of b given attributes.

For each sensor system k, at time t_j, B_{jk} is known r_{jk} by m geolocation target measurement matrix, and R_{jk} is a known corresponding r_{jk} by r_{jk} positive definite measurement error covariance matrix. $m \equiv \dim(X_i)$ is known; X_i is unknown i^{th} target state parameter vector. See Eqs. (2.1) and (2.2) for further explanations and other related definitions.

Prior to t_j, partitioning $Q^{(j-1)}$, and target state parameter vectors, one-step predictions from the Kalman filter are available: $\hat{X}_{i;j-1,j}$, state estimator; and $\Lambda_{i,j-1,j}$, covariance matrix of extimator error; $i = 1, 2, ..$

Also, available prior to t_j are geolocation scores $L_i^{(j-1)}$ $i = 0,1,..$ ($i = 0$, corresponding to false alarm set: $i \geq 1$ corresponding to i^{th} target track set, determined by $Q^{(j)}$), non-geolocation attribute scores $\mathcal{L}_i^{(j-1)}$, $i = 1,2,...,$ possible Kalman filter predictions for H_i - true b by 1 attribute vector of target i - as $\hat{H}_{i;j-1,j}$, $\hat{\Lambda}_{i;j-1,j}$, etc.

Also available prior to t_j are the overall geolocation target data score, $-2 \log pr(Z_+^{(j-1)}|Q^{(j-1)})$, the overall false alarm (geolocation) data score, $-2 \log pr(Z_0^{(j-1)}|Q^{(j-1)}) = L_0^{(j-1)}$, and the overall non-geolocation target attribute score $-2 \log pr(Y^{(j-1)}|Q^{(j-1)})$.

Following, the reception of new data Z_j, based on all of the old data $Z^{(j-1)}$ and the new, combined into $Z^{(j)}$, and based on possibly old partitioning $Q^{(j-1)}$, new partitioning $Q^{(j)}$ of $Z^{(j)}$ is determined by the observer.

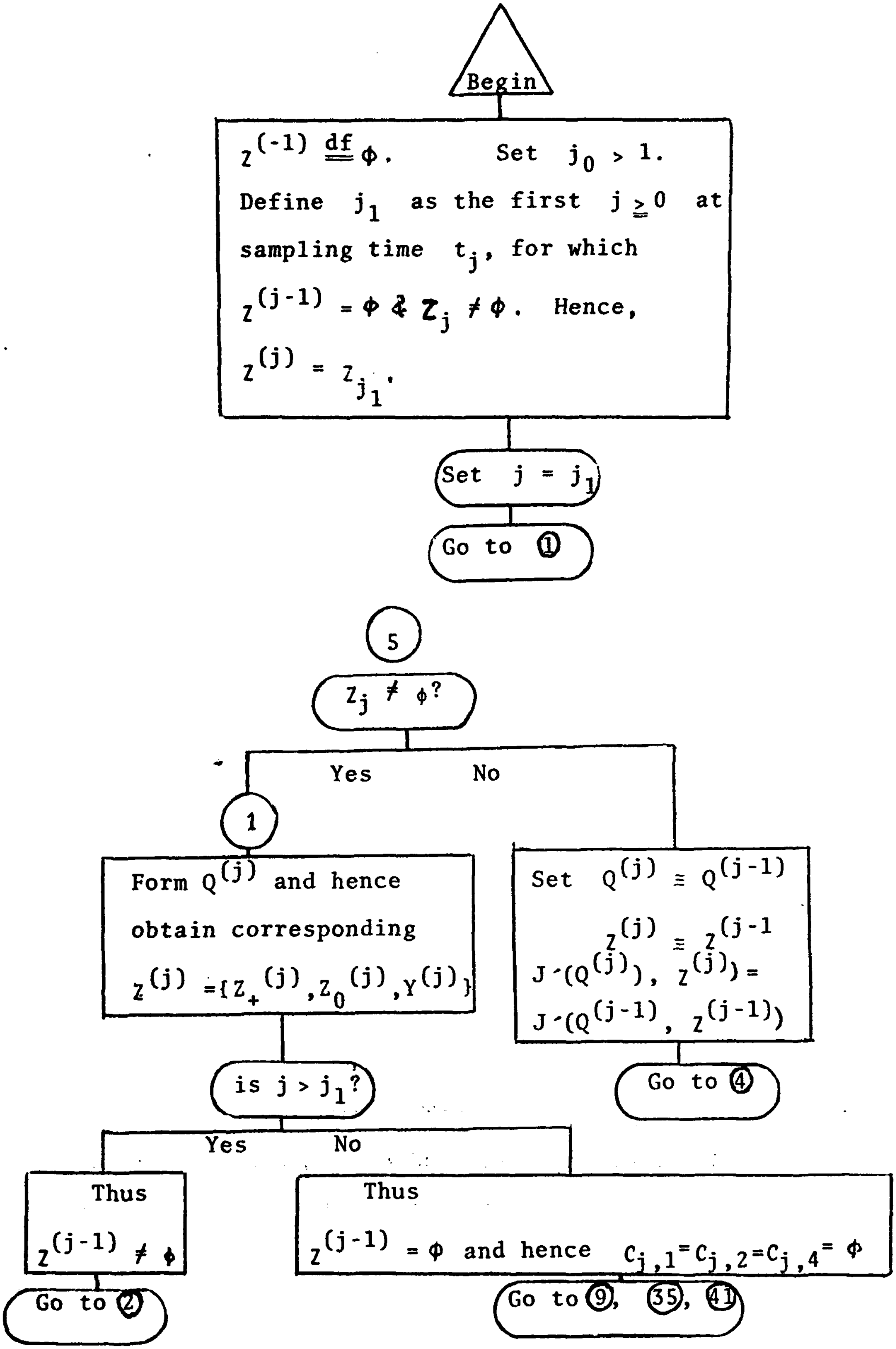
Once $Q^{(j)}$ is determined, then the total data $Z^{(j)}$ can be decomposed as $Q^{(j)} = \{Q_i^{(j)}|i = 0, 1,2,...\}$, where the i^{th} track set (target, for $i \geq 1$, false alarms for $i = 0$) is $Q_i^{(j)} = Z_i^{(j)} = \{Z_{i\alpha k} \mid 0 \leq \alpha \leq j,\ 1 \leq k \leq q\}$. The $Z_{i\alpha k}$'s are the same as the $Z_{\gamma\alpha k}$'s with the i's and γ's replacing each other, except for $Z_{0\alpha k} = Z_{0\alpha k} = \{Z_{0\alpha k_\omega} \mid 1 \leq \omega \leq f_{jk}\}$, the false alarm set, where $f_{jk} \geq 0$ known relative to $Q^{(j)}$. (All $Z_{0\alpha k_\omega}$'s, as well as all $Z_{i\alpha k}$'s, unless vacuous are, r_{jk} by 1.)

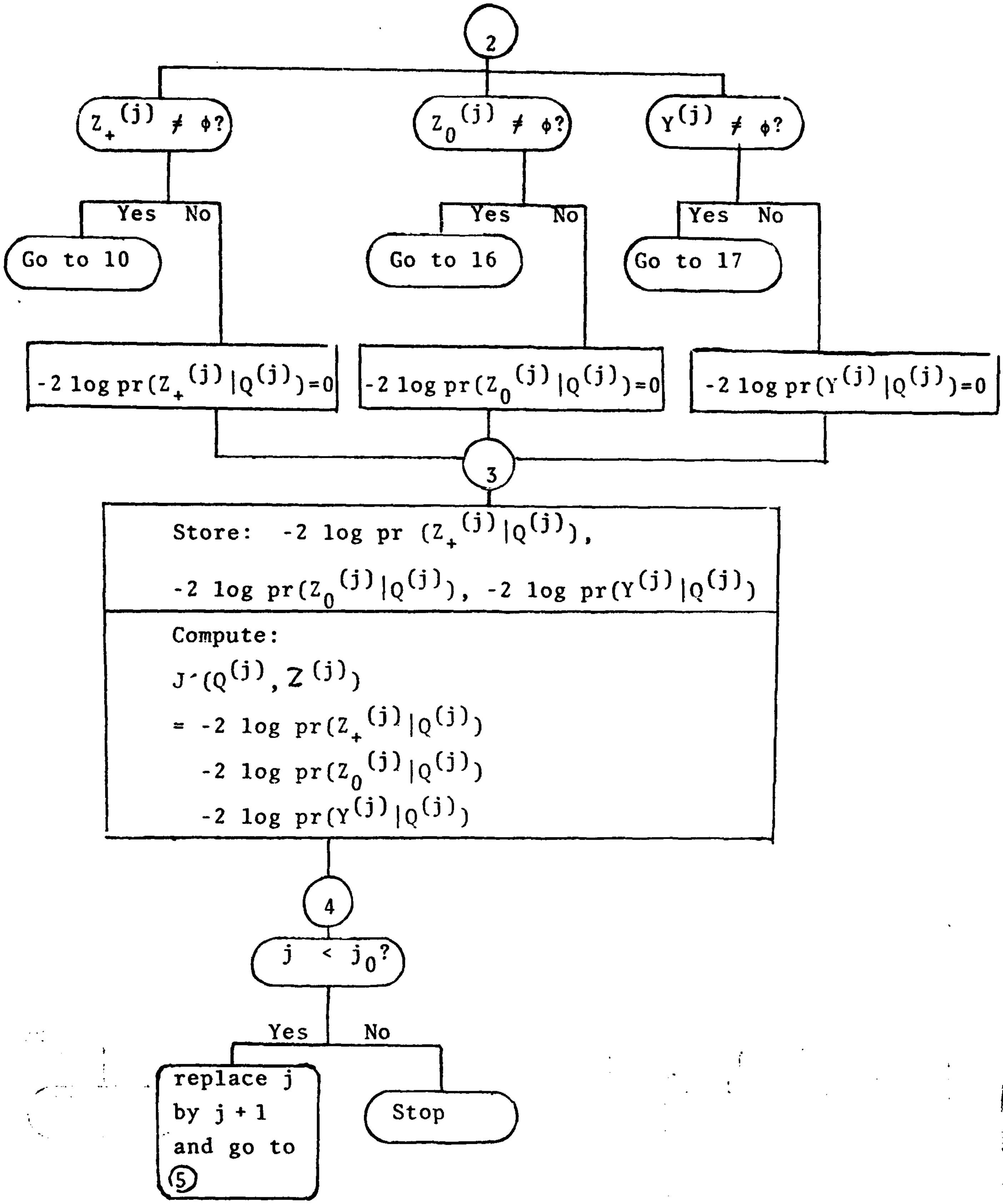
Some of the $Z_{i\alpha k}$'s may be vacuous ($Z_{i\alpha k} = \phi$), in which case, those are generally marked so, deleted from the set, and do not contribute to the computations (actually contribute zero in value to the various sums calculated). Relative to $Q^{(j)}$, i is <u>formally</u> assumed known and $Z_{i\alpha k} = \begin{pmatrix} Z_{i\alpha k} \\ \hdashline Y_{i\alpha k} \end{pmatrix}$ $Z_{i\alpha k}$ is the $r_{\alpha k}$ by 1 geolocation data component and $Y_{i\alpha k}$ is the non-geolocation target attribute data component - a_{ijk} by 1. Furthermore, random set $T_{ijk} = \{b_{ijk,1}, \ldots, b_{ijk;a_{ijk}}\} \subseteq \{1, \ldots, b\}$ is observed, and may be vacuous, in which case $a_{ijk} = 0$. Without loss of generality, $1 \leqq b_{ijk,1} < \cdot < b_{ijka_{ijk}} \leqq b$. Associated with the $Y_{i\alpha k}$'s are measurement matrix $\mathcal{B}_{i\alpha k}$ and error covariance matrix $\mathcal{R}_{i\alpha k}$ (see remarks between eqs. (4.70) and (4.71), and eq. (2.3) and following discussion.)

Note also the notation: $Z_0^{(j)} \equiv Z_0^{(j)} = \{Z_{0\alpha k} | 0 \leqq \alpha \leqq j\}$; false alarm set up to t_j; $Z_+^{(j)} = \{Z_i^{(j)} \mid i \geqq 1\}$, for the set of all geolocation data $Z_i^{(j)} = \{Z_{i\alpha k} | 1 \leqq k \leqq q;\ 0 \leqq \alpha \leqq j\ (Z_{i\alpha k} \neq \phi)\}$ of track set i; up to t_j, for all $i \geqq 1$, and $Y^{(j)} = \{Y_i^{(j)} | i \geqq 1\}$, for the set of all non-geolocation data of track set i, $Y_i^{(j)} = \{Y_{i\alpha k} | 1 \leqq k \leqq q;\ 0 \leqq \alpha \leqq j\ (Y_{i\alpha k} \neq \phi)\}$, up to t_j, for all $i \geqq 1$.

For additional clarifications and definitions, see the main text, especially subsection 2 of the Analysis Section.

The general convention assumed in these flow charts is that any sum whose index set of summation is the empty set (ϕ), is set equal to zero.


Begin
$Z^{(-1)} \stackrel{df}{=} \phi$. Set $j_0 > 1$.
Define j_1 as the first $j \geq 0$ at sampling time t_j, for which $Z^{(j-1)} = \phi$ & $Z_j \neq \phi$. Hence, $Z^{(j)} = Z_{j_1}$.
Set $j = j_1$
Go to 1
5
$Z_j \neq \phi$?
Yes
No
1
Form $Q^{(j)}$ and hence obtain corresponding $Z^{(j)} = \{Z_+{}^{(j)}, Z_0{}^{(j)}, Y^{(j)}\}$
Set $Q^{(j)} \equiv Q^{(j-1)}$
$Z^{(j)} \equiv Z^{(j-1}$
$J'(Q^{(j)}, Z^{(j)}) = J'(Q^{(j-1)}, Z^{(j-1)})$
Go to 4
is $j > j_1$?
Yes
No
Thus $Z^{(j-1)} \neq \phi$
Go to 2
Thus $Z^{(j-1)} = \phi$ and hence $C_{j,1} = C_{j,2} = C_{j,4} = \phi$
Go to 9, 35, 41


2
$Z_+^{(j)} \neq \phi$?
$Z_0^{(j)} \neq \phi$?
$Y^{(j)} \neq \phi$?
Yes
No
Go to 10
Go to 16
Go to 17
$-2 \log pr(Z_+^{(j)}|Q^{(j)})=0$
$-2 \log pr(Z_0^{(j)}|Q^{(j)})=0$
$-2 \log pr(Y^{(j)}|Q^{(j)})=0$
3
Store: $-2 \log pr(Z_+^{(j)}|Q^{(j)})$, $-2 \log pr(Z_0^{(j)}|Q^{(j)})$, $-2 \log pr(Y^{(j)}|Q^{(j)})$
Compute:
$J'(Q^{(j)}, Z^{(j)})$
$= -2 \log pr(Z_+^{(j)}|Q^{(j)})$
$-2 \log pr(Z_0^{(j)}|Q^{(j)})$
$-2 \log pr(Y^{(j)}|Q^{(j)})$
4
$j < j_0$?
Yes
No
replace j by j + 1 and go to 5
Stop

10

$C_{j4} \stackrel{df}{=} \{i \mid 1 \leq i \ \&\ Z_i^{(j)} \neq \phi \ \&\ i \notin C_{ij1} \cup C_{ij2} \cup C_{ij3}\}$

= set of all track sets consisting of two or more points which are obtained by breaking up previous adjudged false alarms, given $Q^{(j)} \neq \phi$?

This is equivalent to: $Z_+^{(j-1)} \not\subset Z_+^{(j)}$?,

i.e. are some previous track sets broken up?

Yes

Use old outouts $L_i^{(j-1)}$ for $i \in C_{j1} \cup C_{j2} \cup C_{j3} \cup C_{j4}$

Go to 7, 8, 9, 15

No

Note that $C_{j4} = \phi$

Go to 7, 8, 9

6

$$-2 \log pr\,(Z_+^{(j)} \mid Q^{(j)})$$

$$= -2 \log pr\,(Z_+^{(j-1)} \mid Q^{(j-1)}) + \sum_{(i \in C_{j1} \cup C_{j2} \cup C_{j3})} L_{ij}$$

Also store all $L_i^{(j)}$'s and Kalman filter outputs,

Go to 3

7

$C_{j1} \stackrel{df}{=} \{i \le l \mid i \;\&\; \phi \ne Z_i^{(j-1)} \subset Z_i^{(j)} \; (Z_{ij} \ne \phi)\}$

= set of previous track sets which will be nontrivially updated by new data, given $Q^{(j)}$

$\ne \phi$?

Yes: Go to 11

No: $C_{j4} \ne \phi$?

Yes: Set $C_{j1} = \phi$. Go to 12

No: Set $C_{j1} = \phi$ and $C_{j4} = \phi$ Go to 6

12

$-2 \log pr(Z_+^{(j)} \mid Q^{(j)})$

$= \sum L_i^{(j)}$

$(i \epsilon C_{j1} \cup C_{j2} \cup C_{j3} \cup C_{j4})$.

Also store all $L_i^{(j)}$'s and Kalman filter outputs

Go to 3

8

$C_{j2} \stackrel{df}{=} \{i \mid 1 \leq i \ \& \ Z_i^{(j-1)} = Z_i^{(j)} \neq \phi \text{ (and } Z_{ij} = \phi)\}$

= set of all previous track sets which will be updated by predictions only (no new data), given $Q^{(j)}$

$\neq \phi$?

Yes: Go to 13

No: $C_{i4} \neq \phi$?

- Yes: Set $C_{j2} = \phi$; Go to 12
- No: Set $C_{j2} = \phi$ and $C_{j4} = \phi$; Go to 6

9

$C_{j3} \stackrel{df}{=} \{i \mid 1 \leq i \ \& \ Z_i^{(j)} = Z_{ij} \neq \phi \text{ (and } Z_i^{(j-1)} = \phi)\}$

= set of all new (one point) track sets, given $Q^{(j)}$

$\neq \phi$?

Yes: Go to 14

No: $C_{j4} \neq \phi$?

- Yes: Set C_{j3} =; Go to 12
- No: Set $C_{j3} = \phi$ and $C_{j4} = \phi$; Go to 6

11

Inputs: For $i \in C_{j1}$

From before: $\hat{X}_{i;j-1,j}$, $\Lambda_{i;j-1,j}$, $s_i^{(j-1)} = \sum_{\alpha=0}^{j-1} s_{i\alpha}$, $L_i^{(j-1)''}$, $L_i^{(j-1)}$

At present: r_{jk} known positive number. B_{jk} (r_{jk} by m) for k such that $Z_{ijk} \neq \phi$

R_{jk} (r_{jk} by r_{jk} positive definite), for k such that $Z_{ijk} \neq \phi$

$\{Z_{ijk} \mid i \in C_{j1};\ 1 \leq k \leq q$ such that $Z_{ijk} \neq \phi\}$

$$s_{ij} = \sum_{\substack{1 \leq k \leq q \\ \text{such that} \\ Z_{ijk} \neq \phi}} r_{jk}, \text{ for } i \in C_{j1}$$

$$s_i^{(j)} = s_i^{(j-1)} + s_{ij}$$

$$\nu_{ijk} = Z_{ijk} - B_{jk} \cdot \hat{X}_{ij-1,j};\ i \in C_{j1}$$

$s_{ij} > m$?

No: Go to 24

Yes: 25

$$D_{ij} = \sum_{\substack{1 \leq k \leq q \\ \text{such that} \\ Z_{ijk} \neq \phi}} B_{jk}^T R_{jk}^{-1} B_{jk}$$

$$\Lambda_{ijj} = (D_{ij} + \Lambda_{i,j-1,j}^{-1})^{-1}$$

$$= D_{ij}^{-1} \cdot (I_m - (I_m + D_{ij} \Lambda_{i,j-1,j})^{-1})$$

$$A_{ijk'k''} = \delta_{k'k''} \cdot R_{jk'}^{-1} - R_{jk'}^{-1} B_{jk'} \Lambda_{ijj} B_{jk''}^T R_{jk''}^{-1}$$

for $1 \leq k', k'' \leq q$ such that $Z_{ijk'} \neq \phi$, $Z_{ijk''} \neq \phi$

$$\log \det \Sigma_{ij} = \sum_{\substack{1 \leq k \leq q \\ \text{such that} \\ Z_{ijk} \neq \phi}} \log \det R_{jk}$$

$$+ \log \det (I_m + D_{ij} \cdot \Lambda_{i,j-1,j})$$

Go to 26

13

Inputs for $i \varepsilon C_{j2}$

From before: $\hat{X}_{i,j-1,j}, \Lambda_{i;j-1,j}, s_i^{(j-1)} = \sum_{\alpha=0}^{j-1} s_{i\alpha}, L_i^{(j-1)}$

At present: No inputs

(No r_{jk}'s, B_{jk}'s, $\dot{R}_{jk}$'s, Z_{ijk}'s)

$s_i^{(j)} = s_i^{(j-1)}$

$\hat{X}_{i;jj} = \hat{X}_{i;j-1,j}$

$\Lambda_{i;j,j} = \Lambda_{i;j-1,j}$

$L_{ij} = 0$

$L_i^{(j)} = L_i^{(j-1)}$

Go to 27

28

$C_{j4} \neq \phi$?

Yes: Go to 12

No: Go to 6

24

$$G_{ijk',k''} = \delta_{k'k''} R_{jk'} + B_{jk'} \Lambda_{i,j-1,j} B^{T}_{jk''}$$

for $1 \le k', k'' \le q$, such that $Z_{ijk'} \ne \phi$, $Z_{ijk''} \ne \phi$

For $\Sigma_{ij} = (G_{ijk',k''})_{\substack{1 \le k',k'' \le q \\ Z_{ijk'} \ne \phi \\ Z_{ijk''} \ne \phi}}$,

compute

log det (Σ_{ij}) and

$$\Sigma^{-1}_{ij} \overset{df}{=} (A_{ijk',k''})_{\substack{1 \le k',k'' \le q \\ \text{such that} \\ Z_{ijk'} \ne \phi,\ Z_{ijk''} \ne \phi}}$$

Appendix A may be useful here

Go to 26

29

$$F_{ij} = \sum_{\substack{1 \le k',k'' \le q, \\ \text{such that} \\ Z_{ijk'} \ne \phi, \\ Z_{ijk''} \ne \phi}} B^{T}_{jk'} A_{ijk',k''} B_{jk''}$$

$$\Lambda_{ijj} = (I_m - \Lambda_{i,j-1,j} \cdot F_{ij}) \Lambda_{i,j-1,j}$$

Go to 27

(26)

$$L'_{ij} = s_{ij} \log 2\pi + \log \det \Sigma_{ij}$$

$$L''_{ij} = \sum_{\substack{1 \le k', k'' \le q \\ \text{such that} \\ Z_{ijk'} \ne \phi, \\ Z_{ijk''} \ne \phi}} \nu^T_{ijk'} \cdot A_{ijk'k''} \cdot \nu_{ijk''}$$

$$L_{ij} = L'_{ij} + L''_{ij}$$

$i \in C_{j3}$ → Go to (30)

$i \in C_{j1}$

$$L_i^{(j)} = L_i^{(j-1)} + L_{ij}$$

(31) →

$$L_i^{(j)''} = L_i^{(j-1)''} + L''_{ij}$$

For each $i \in C_{j1}$

$L_i^{(j)''} > T_{ij}(\beta)$, ?

where T_{ij} is determined from

$\beta = \Pr(\chi^2_{s_i}(j) > T_{ij}(\beta))$

for $0 < \beta < 1$, β small

No: Continue track set i. Redefine C_{j1} to be restricted to i such that $L_i^{(j)''} \le T_{ij}(\beta)$.

Yes: M is a fixed large number. Begin formally new track set: replace i by $M+i$. Put $M+i \in C_{j3}$. Set $Z_{M+1}^{(j-1)j3} = \phi$.

Go to (9)

(32) → $s_{ij} > m$?

Yes: Go to (33)

No: Go to (34)

33

$$\hat{X}_{ijj} = \hat{X}_{i,j-1,j} + \Lambda_{ijj} \cdot \sum_{\substack{1 \leq k \leq q \\ \text{such that} \\ Z_{ijk} \neq \phi}} B^T_{jk} \cdot R^{-1}_{jk} \cdot \nu_{ijk}$$

Go to 27

34

$$\hat{X}_{ijj} = \hat{X}_{ij-1,j} + \Lambda_{ij-1,} \sum_{\substack{1 \leq k',k'' \leq q \\ \text{such that} \\ Z_{ijk'} \neq \phi \\ Z_{ijk''} \neq \phi}} B^T_{jk'} \cdot A_{ijk'k''} \nu_{ijk''}$$

Go to 29

27

$$\hat{X}_{i,j,j+1} = \Phi_{j+1} \cdot \hat{X}_{ijj} \quad ,$$

$$\Lambda_{i,j,j+1} = \Phi_{j+1} \cdot \Lambda_{i,jj} \cdot \Phi^T_{jk} + G_{ij} \cdot P_j \cdot G^T_{ij}$$

Go to 28

Inputs: $i \in C_{j3}$

From before: $L_i^{(j-1)} = 0$

Formally set at present: $\hat{X}_{i,j-1,j} = E(X_{ij}), \Lambda_{i,j-1,j} = \mathrm{Cov}(X_{ij})$

At present: r_{jk}, B_{jk} (r_{jk} by m) for k such that $Z_{ijk} \neq \phi$,

R_{jk} (r_{jk} by r_{jk} positive definite) for k

such that $Z_{ijk} \neq \phi$,

$\{Z_{ijk} \mid i \in C_{j2};\ 1 \leq k \leq q$ such that $Z_{ijk} \neq \phi\}$

$$s_{ij} = \sum_{\substack{1 \leq k \leq q \\ \text{such that} \\ Z_{ijk} \neq \phi}} r_{jk}$$

$$\nu_{ijk} = Z_{ijk} - B_{jk} E(X_{i,j})$$

$s_{ij} > m$?

Yes: Go to 25

No: Go to 24

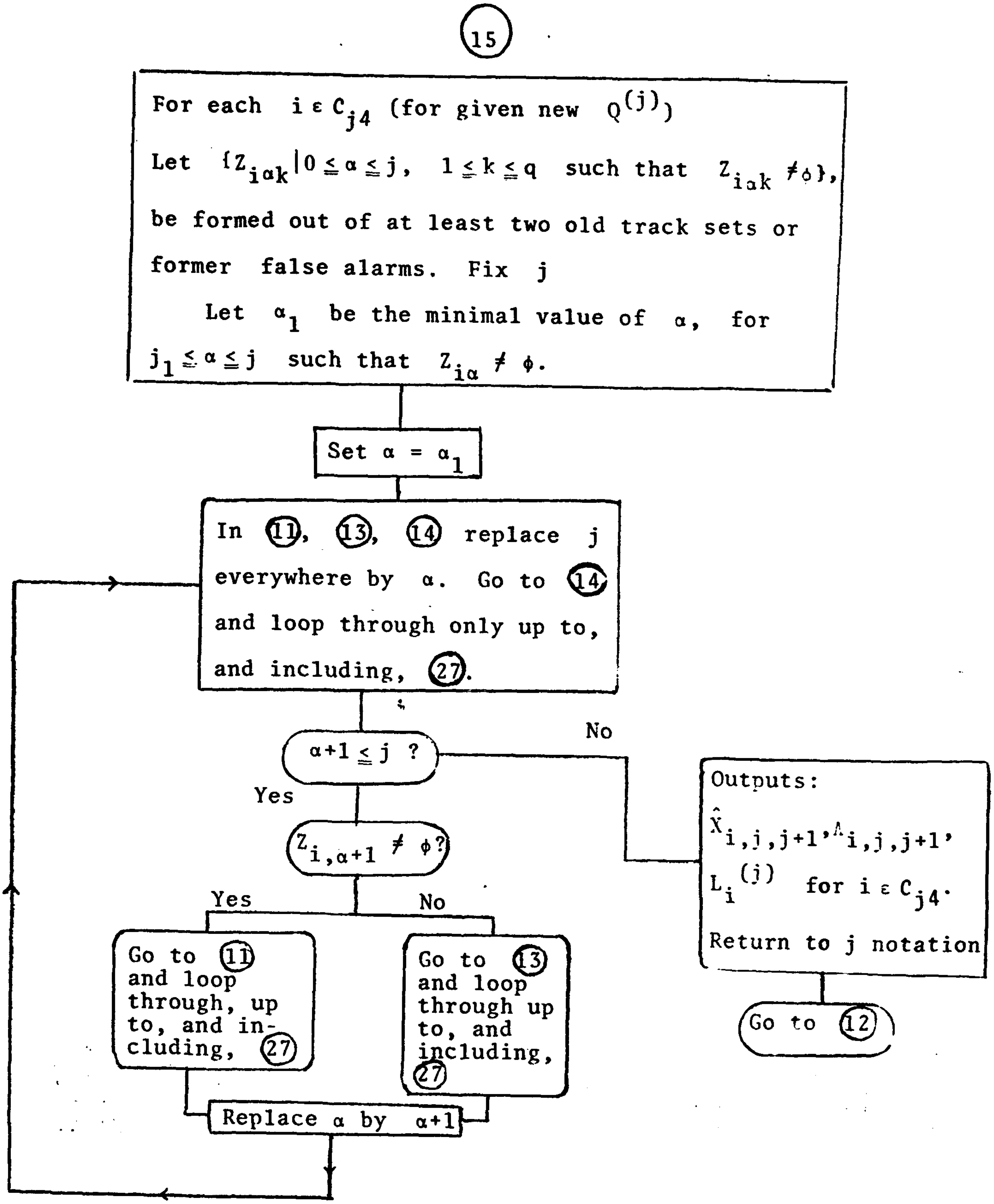
30

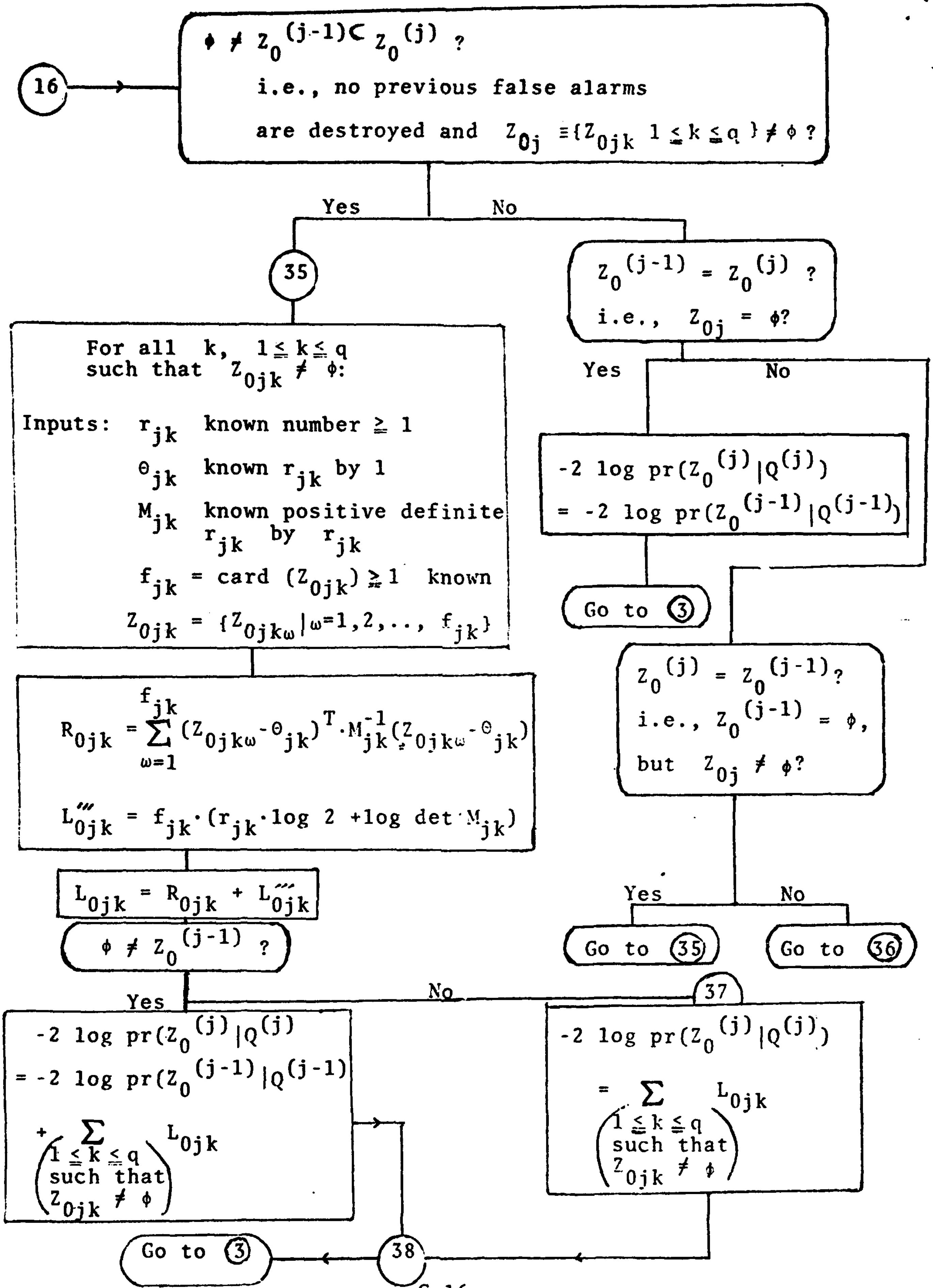
For each $i \in C_{j2}$:
$L_i^{(j)} = L_{ij}$

$s_{ij} > m$?

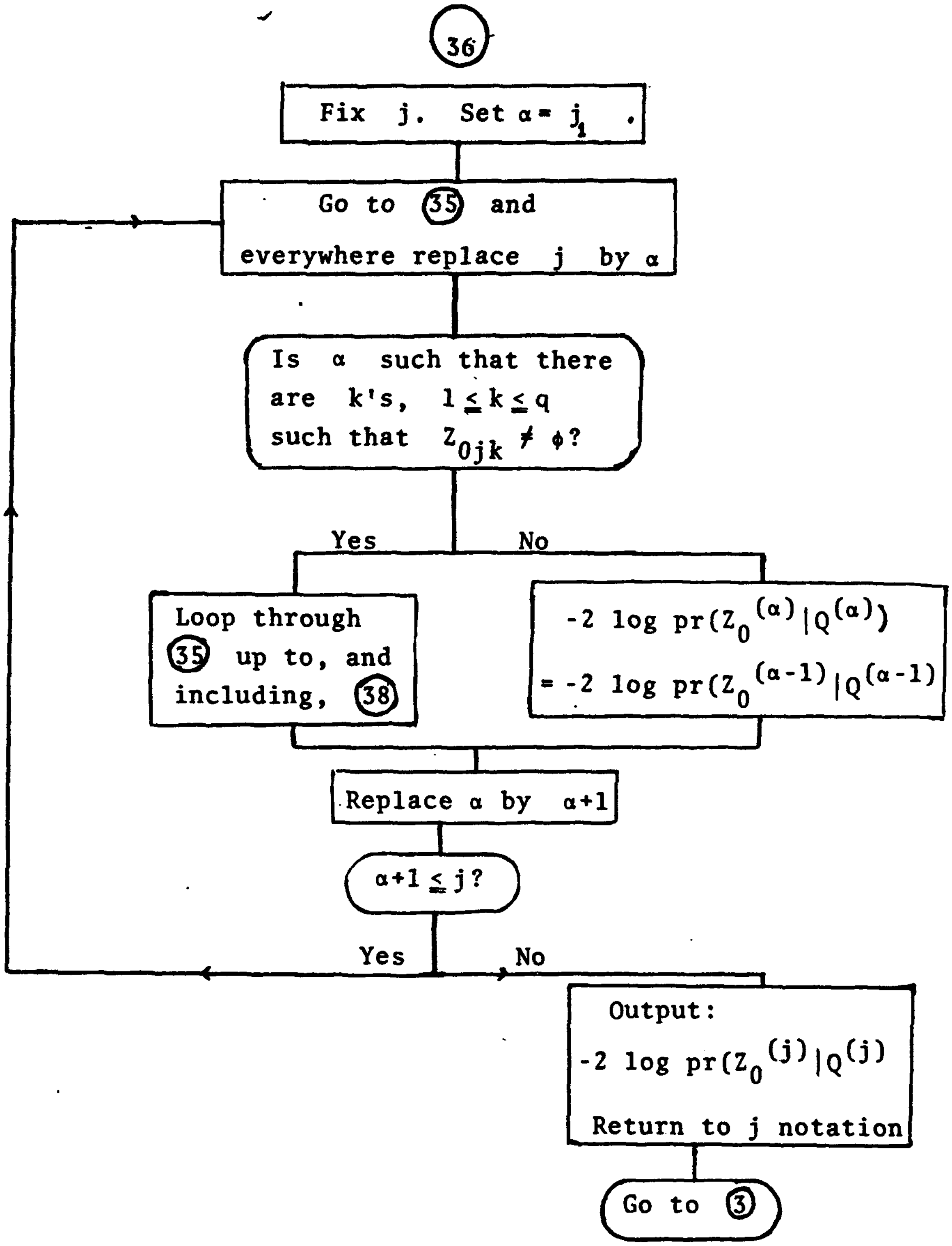
Yes: Go to 33

No: Go to 34

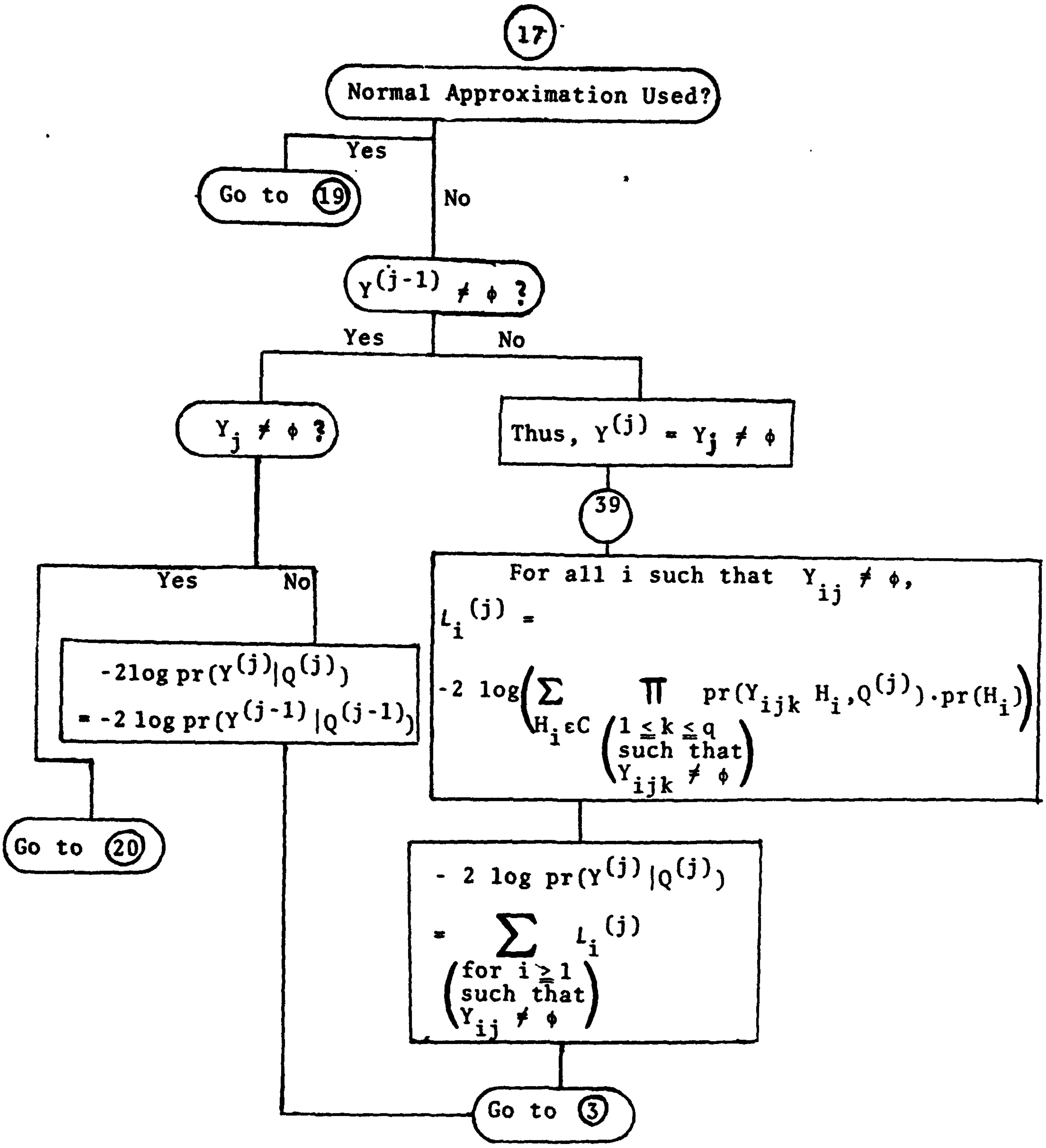

15
For each $i \varepsilon C_{j4}$ (for given new $Q^{(j)}$)
Let $\{Z_{i\alpha k} | 0 \leqq \alpha \leqq j,\ 1 \leqq k \leqq q$ such that $Z_{i\alpha k} \neq \phi\}$, be formed out of at least two old track sets or former false alarms. Fix j
Let α_1 be the minimal value of α, for $j_1 \leq \alpha \leq j$ such that $Z_{i\alpha} \neq \phi$.
Set $\alpha = \alpha_1$
In 11, 13, 14 replace j everywhere by α. Go to 14 and loop through only up to, and including, 27.
$\alpha+1 \leq j$?
No
Yes
$Z_{i,\alpha+1} \neq \phi$?
Yes
No
Go to 11 and loop through, up to, and including, 27
Go to 13 and loop through up to, and including, 27
Replace α by $\alpha+1$
Outputs:
$\hat{X}_{i,j,j+1}, \Lambda_{i,j,j+1}$,
$L_i^{(j)}$ for $i \varepsilon C_{j4}$.
Return to j notation
Go to 12

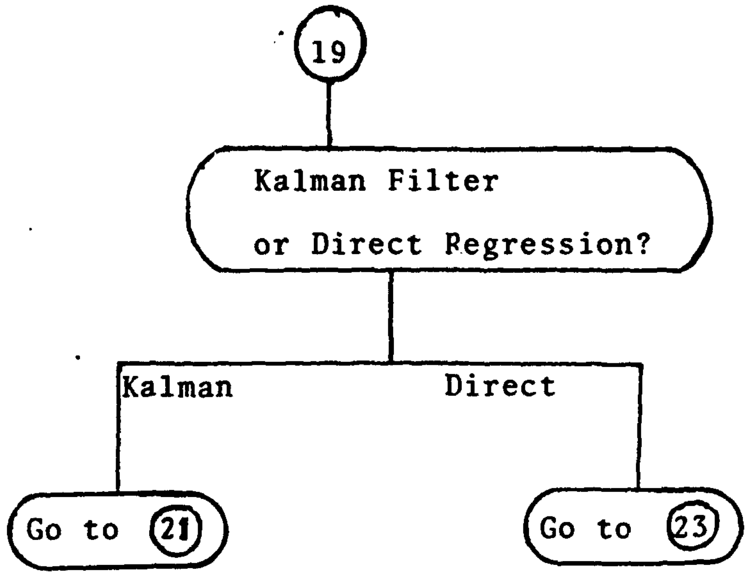


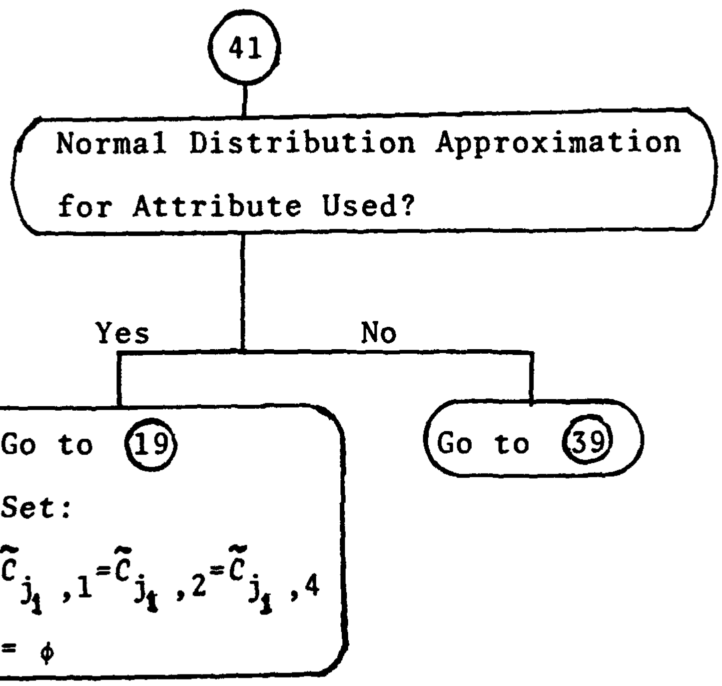
16
$\phi \neq Z_0^{(j-1)} \subset Z_0^{(j)}$?
i.e., no previous false alarms
are destroyed and $Z_{0j} \equiv \{Z_{0jk}\ 1 \leq k \leq q\} \neq \phi$?
Yes
No
35
For all k, $1 \leq k \leq q$
such that $Z_{0jk} \neq \phi$:
Inputs: r_{jk} known number ≥ 1
Θ_{jk} known r_{jk} by 1
M_{jk} known positive definite r_{jk} by r_{jk}
$f_{jk} = \text{card}\ (Z_{0jk}) \geq 1$ known
$Z_{0jk} = \{Z_{0jk\omega} \mid \omega = 1, 2, \ldots, f_{jk}\}$
$R_{0jk} = \sum_{\omega=1}^{f_{jk}} (Z_{0jk\omega} - \Theta_{jk})^T \cdot M_{jk}^{-1} (Z_{0jk\omega} - \Theta_{jk})$
$L'''_{0jk} = f_{jk} \cdot (r_{jk} \cdot \log 2 + \log \det M_{jk})$
$L_{0jk} = R_{0jk} + L'''_{0jk}$
$\phi \neq Z_0^{(j-1)}$?
Yes
No
$-2 \log pr(Z_0^{(j)} | Q^{(j)})$
$= -2 \log pr(Z_0^{(j-1)} | Q^{(j-1)})$
$+ \sum_{\substack{1 \leq k \leq q \\ \text{such that} \\ Z_{0jk} \neq \phi}} L_{0jk}$
Go to 3
38
37
$-2 \log pr(Z_0^{(j)} | Q^{(j)})$
$= \sum_{\substack{1 \leq k \leq q \\ \text{such that} \\ Z_{0jk} \neq \phi}} L_{0jk}$
$Z_0^{(j-1)} = Z_0^{(j)}$?
i.e., $Z_{0j} = \phi$?
Yes
No
$-2 \log pr(Z_0^{(j)} | Q^{(j)})$
$= -2 \log pr(Z_0^{(j-1)} | Q^{(j-1)})$
Go to 3
$Z_0^{(j)} = Z_0^{(j-1)}$?
i.e., $Z_0^{(j-1)} = \phi$,
but $Z_{0j} \neq \phi$?
Yes
No
Go to 35
Go to 36


36
Fix j. Set α = j₁ .
Go to 35 and everywhere replace j by α
Is α such that there are k's, 1 ≤ k ≤ q such that $Z_{0jk} \neq \phi$?
Yes
No
Loop through 35 up to, and including, 38
$-2 \log pr(Z_0^{(\alpha)} | Q^{(\alpha)})$
$= -2 \log pr(Z_0^{(\alpha-1)} | Q^{(\alpha-1)})$
Replace α by α+1
α+1 ≤ j?
Yes
No
Output:
$-2 \log pr(Z_0^{(j)} | Q^{(j)})$
Return to j notation
Go to 3

Non-Geolocation Target Attribute Data




19
Kalman Filter or Direct Regression?
Kalman
Direct
Go to 21
Go to 23


41
Normal Distribution Approximation for Attribute Used?
Yes
No
Go to 19
Set:
$\tilde{C}_{j_1,1}=\tilde{C}_{j_1,2}=\tilde{C}_{j_1,4} = \phi$
Go to 39

20

For all $i \geq 1$ such that $Y_i^{(j)} \neq \phi$; $H_i \in C$,

$$pr(Y_i^{(j)} | H_i, Q^{(j)}) = \prod_{\alpha=0}^{j} \prod_{\begin{pmatrix}1 \leq k \leq q \\ \text{such that} \\ Y_{i\alpha k} \neq \phi\end{pmatrix}} pr(Y_{i\alpha k} | H_i, Q^{(j)})$$

$$L_i^{(j)} = -2 \log \sum_{H_i \in C} pr(Y_i^{(j)} | H_i, Q^{(j)}) \cdot pr(H_i)$$

40

$$-2 \log pr(Y^{(j)} | Q^{(j)}) = \sum_{\begin{pmatrix}1 \leq i \\ \text{such that} \\ Y_i(j) \neq \phi\end{pmatrix}} L_i^{(j)}$$

Go to 3

Kalmen Filter Approach for Non-geolocation Target Attribute Data Under the Normal Distributional Approximations

21

Go to 10 and everywhere replace:

m by b	r_{jk} by a_{ijk}
Z_{ijk} by Y_{ijk}	R_{jk} by $\mathcal{R}_{ijk}$
$Z_+^{(j)}$ by $Y^{(j)}$	B_{jk} by $\mathcal{B}_{ijk}$
$Z_i^{(j)}$ by $Y_i^{(j)}$	ϕ_j by I_b
etc.	G_j by 0
L by $\mathcal{L}$	P_j by 0
$L_i^{(j)}$ by $\mathcal{L}_i^{(j)}$	(Thus 27 becomes
L_{ij} by $\mathcal{L}_{ij}$	$\hat{H}_{ij,j+1} = \hat{H}_{ijj}$
	$\tilde{\Lambda}_{ij,j+1} = \tilde{\Lambda}_{ijj}$.)
C_{jl} by $\tilde{C}_{jl}$	D_{ij} by $\mathcal{D}_{ij}$
etc.	
X_{ij} by H_i,	$A_{ijk',k''}$ by $\mathcal{A}_{ijk'k''}$
$\hat{X}_{i;j-1,j}$ by $\hat{H}_{i;j-1,j}$	Σ_{ij} by $\tilde{\Sigma}_{ij}$
$\hat{X}_{i;j,j}$ by $\hat{H}_{i;j,j}$	$G_{ijk'k''}$ by $\mathcal{G}_{ijk'k''}$
etc.	
$\Lambda_{i;j-1,j}$ by $\tilde{\Lambda}_{i;j-1,j}$	F_{ij} by $\mathcal{F}_{ij}$
etc.	

Go to 22

22

Also everywhere in 10 replace:

$s_i^{(j-1)}$ by $\delta_i^{(j-1)}$

s_{ij} by δ_{ij}

ν_{ijk} by $\tilde{\nu}_{ijk}$.

Further evaluations:

For $\tilde{\nu}_{ijk}$, use (4.98).

If $\delta_{ij} > b$:

For $\mathcal{D}_{ij}$, use eqs. (4.93) - (4.95).

For $A_{ijk'k''}$, use (4.91), (4.101), (4.102).

For log det (R_{ijk}), use (4.103).

For $B_{ijk}^T R_{ijk}^{-1} \cdot \tilde{\nu}_{ijk}$, use (4.96), (4.97).

For $\delta_{ij} \leq b$:

For $B_{ijk'} \tilde{\Lambda}_{i,j-1,j} B_{ijk''}^T$, use analogue of eq. (4.102).

For $B_{ijk'}^T \cdot A_{ijk'k''} \cdot B_{ijk''}$, use eqs. (4.109) - (4.111).

For $B_{ijk'}^T \cdot A_{ijk'k''} \cdot \tilde{\nu}_{ijk''}$, use eqs. (4.98), (4.109), (4.112), (4.113).

When 31 is reached, skip to 32 and continue looping through as usual.

When 3 is reached and outputs deposited, resume old notation again in 10, until next cycle into 19.

Direct Regression Approach to Non-geolocation Target Attribute Data , Under Normal Distributional Approximations

23

$$s_i^{(j)} = \sum_{\alpha=j_1}^{j} s_{i\alpha}$$

$s_i^{(j)} > b?$

No: Go to 21

Yes:

$$\log \det R_i^{(j)} = \sum_{\alpha=j}^{j} \log \det (R_{i\alpha})$$

$$\log \det R_{i\alpha} = \sum_{\substack{1 \le k \le q \\ \text{such that} \\ Y_{i\alpha k} \ne \phi}} \log \det R_{i\alpha k}$$

$$\log \det R_{i\alpha k} = \sum_{\gamma=1}^{a_{i\alpha k}} \log r_{b_{i\alpha k\gamma}}$$

$$d_{i,\gamma}^{(j)} = \sum_{\alpha=j_1}^{j} d_{i\alpha\gamma} = \left(1/r_\gamma\right) \cdot \sum_{\substack{1 \le k \le q,\ j_1 \le \alpha \le j, \\ \text{such that } Y_{i\alpha k} \ne \phi, \\ \text{and there is } \eta,\ 1 \le \eta \le b \\ \text{such that } \gamma = b_{i\alpha k\eta} \\ \in T_{i\alpha k}}} (1)$$

$$\log \det (I_b + D_i^{(j)} \cdot \mathrm{Cov}(H_i))$$

$$= \sum_{\gamma=1}^{b} \log(1 + d_{i\gamma}^{(j)} \cdot \mathrm{var}(H_{i\gamma}))$$

$$\log \det \mathrm{Cov}(Y_i^{(j)} | Q^{(j)})$$

$$= \log \det R_i^{(j)} + \log \det (I_b + D^{(j)} \cdot \mathrm{Cov}(H_i))$$

Go to 18

18

$$L_i^{(j)\prime} = \delta_i^{(j)} \log 2\pi + \log \det \mathrm{Cov}(Y_i^{(j)} \mid Q^{(j)})$$

$$g_{i,\gamma}^{(j)} = \left(1/r_\gamma\right) \cdot \sum_{\substack{j_1 \le \alpha \le j}} \; \sum_{\substack{1 \le k \le q}} \left(Y_{i\alpha k} - E(H_{i,b_{i\alpha k\eta}})\right)$$

$\left(\begin{array}{l} j_1 \le \alpha \le j \quad 1 \le k \le q \\ \text{such that } Y_{i\alpha k} \ne \phi \\ \text{and such that there} \\ \text{is } \eta,\ 1 \le \eta \le b, \text{ so that} \\ \gamma = b_{i\alpha k\eta} \in T_{i\alpha k} \end{array} \right)$

Approximation desired minimizing knowledge of $E(H_i)$, $\mathrm{Cov}(H_i)$?

Yes — No

In computation for $L_i^{(j)''}$ below, set $E(H_{i,b_{i\alpha k\eta}}) = 0$ and $\mathrm{var}(H_{i\gamma}) = +\infty$

$$L_i^{(j)''} = \sum_{\alpha=j_1}^{j} \; \sum_{\substack{1 \le k \le q \\ \text{such that} \\ Y_{i\alpha k} \ne \phi}} \; \sum_{\eta=1}^{a_{i\alpha k}} \left(1/r_{b_{i\alpha k\eta}}\right) \cdot \left(Y_{i\alpha k} - E(H_{i,b_{i\alpha k\eta}})\right)^2$$

$$- \sum_{\gamma=1}^{b} \left(1/\left((1/\mathrm{var}(H_{i\gamma})) + d_{i,\gamma}^{(j)}\right)\right) \cdot g_{i,\gamma}^{(j)\,2}$$

$$L_i^{(j)} = L_i^{(j)\prime} + L_i^{(j)''} ;$$

for all i such that $Y_i^{(j)} \ne \phi$.

Go to 40

SUMMARY OF FLOW CHART INPUTS OUTPUTS

Summary of Time-Independent Quantitites

C is a known set of possible non-geolocation target attributes. Card (C) = a. Each element of C is a b by 1 vector of values such as (true) hull length, flag color, identification, etc.

Fixed number of sensor systems operating is q, though some at times may fail to collect data.

m is the common dimension of each target state parameter vector; n is the dimension of the driving noise vector (eq. (2.1)).

If normal approximations are used for the non-geolocation target attributes:

$E(H_i)$, $Cov(H_i)$ are the mean and covariance matrix of variability or randomness for any target i, and it is assumed their values are not dependent on i (nor on j).

If normal approximations are not used for the non-geolocation target attributes:

$pr(H_i)$ the prior probability function of $H_i \varepsilon C$ is known; it is usually assumed to be uniform: $pr(H_i) = {}^1/_a$ for all $H_i \varepsilon C$, where a = card(C). ($pr(H_i)$ does not depend on j.)

Summary of Previous Outputs for Data

Partitioning $Q^{(j-1)} = \{Q_0^{(j-1)}, Q_1^{(j-1)}, \ldots\}$, where $Q_i^{(j-1)}$ is the perceived (target, if $i \geq 1$, false alarm, if $i = 0$) i^{th} track set consisting of data $Z_i^{(j-1)} = \{\mathbf{Z}_{i\alpha k} \mid 0 \leq \alpha \leq j, 1 \leq k \leq q,$ such that $\mathbf{Z}_{i\alpha k} \neq \phi\}$.

Kalman filter outputs: $\hat{X}_{i,j-1,j}$, $\Lambda_{i,j-1,j}$
(m by 1) (m by m positive definite)

$\hat{H}_{i;j-1,j}$, $\tilde{\Lambda}_{i;j-1,j}$. (The last two are not re-
(b by 1) (b by b positive definite)

quired if the Kalman filter - normal approximation procedure is not used for the non-geolocation target attributes.)

$L_i^{(j-1)} = -2 \log pr(Z_i^{(j-1)} | Q^{(j-1)})$, for $i \geq 0$; $L_i^{(j-1)''}$, for $i \geq 1$;

$\mathcal{L}_i^{(j-1)} = -2 \log pr(Y_i^{(j-1)} | Q^{(j-1)})$, for $i \geq 1$;

$$-2 \log pr(Z_+^{(j-1)} | Q^{(j-1)}) = \sum_{i=1,2,..} (L_i^{(j-1)}) ;$$

$$-2 \log pr(Y^{(j-1)} | Q^{(j-1)}) = \sum_{i=1,2,..} (\mathcal{L}_i^{(j-1)}) ;$$

$$\begin{aligned} J'(Q^{(j-1)}, Z^{(j-1)}) = & -2 \log pr(Z_+^{(j-1)} | Q^{(j-1)}) \\ & -2 \log pr(Z_0^{(j-1)} | Q^{(j-1)}) \\ & -2 \log pr(Y^{(j-1)} | Q^{(j-1)}) \end{aligned}$$

Summary of Data Inputs at Sampling Time t_j but Before New Partitioning $Q^{(j)}$ is carried out

Consider sensor system k, k = 1,.., q:

If no new data at all arrives, $Z_j = \phi$

If some new data arrives $Z_j \neq \phi$, and for each sensor k which obtains data, $m_{jk} \geq 1$ is the number of data reports received;

$$Z_j = (Z_{\gamma,j,k})_{\substack{\gamma=1,..,m_{jk} \\ k=1,..,q}}$$

(where above, set $Z_{\gamma jk} \equiv \phi$ for those k's for which $m_{jk} = 0$).

The γ^{th} report received by sensor k at t_j is , $1 \leq \gamma \leq m_{jk}$

$$Z_{\gamma jk} = \begin{pmatrix} Z'_{\gamma jk} \\ Z''_{\gamma jk} \end{pmatrix} ,$$

$Z'_{\gamma jk}$ is r_{jk} by 1 geolocation data vector

$(r_{jk} \geq 1)$

Z''_{jk} is $a_{\gamma jk}$ by 1 non-geolocation data target attribute vector

$(0 \leq a_{\gamma jk} \leq b)$.

It is possible for a given γ, $1 \leq \gamma \leq m_{jk}$, for $m_{jk} \geq 1$, to be such that $Z'_{\gamma jk} = \phi$ or $Z''_{\gamma jk} = \phi$, but not both (otherwise $m_{jk} = 0$).

$Z''_{\gamma jk} = \Phi$ iff $a_{\gamma jk} = 0$.

$Z'_{\gamma jk}$ contains as components, typically measured positions, velocities, etc., from sensor system k at t_j for report γ

$Z''_{\gamma jk}$ contains as components, a given subset $T_{\gamma jk}$ (of size $a_{\gamma jk}$) of a fixed set of b attribute values.

Typically, this could be

$$Z''_{\gamma jk} = \begin{pmatrix} \text{number of radars on-board} \\ \text{hull length} \\ \text{flag color} \end{pmatrix} ,$$

where the <u>total</u> attribute set is {number of radars on-board, identification, shape, hull length, flag color}
($b = 5$ here).

Also given for each sensor k is B_{jk}, r_{jk} by m measurement matrix.

Typically, if the i^{th} target state parameter vector is
($m = 4$)

$$X_{ij} \equiv \begin{pmatrix} \text{x-pos, at } t_j \\ \text{y-pos, at } t_j \\ \text{x-vel. at } t_j \\ \text{y-vel. at } t_j \end{pmatrix} ,$$

and only x- and y- positive measurements can be made at t_j by sensor k, then

$$Z'_{\gamma jk} = \begin{pmatrix} \text{observed x-pos. at } t_j \\ \text{observed y-pos. at } t_j \end{pmatrix}$$

and

$$B_{jk} = \begin{pmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix} \qquad \text{(2 by 4)}$$

($r_{jk} = 2$) . (see eq. (2.2))

Also given is the measurement error covariance matrix.

R_{jk} (r_{jk} by r_{jk} positive definite) is typically obtained approximately as $R_{jk} = \sigma_{jk} \cdot P_{jk} \cdot \sigma_{jk}^T$, where σ_{jk} is a matrix of Jacobians of the transformation from sensor coordinate to cartesian coordinate space, and P_{jk} is the covariance matrix of measurement error of sensor system k at t_j for its natural (sensor) coordinate system. (See also eq. (2.2).)

Also, for each sensor system k, is given θ_{jk} (r_{jk} by 1), the mean of the false alarm dispersion ; M_{jk} the corresponding (r_{jk} by r_{jk}) positive definite covariance matrix.

Also, the common target motion transition matrix Φ_j (m by m) driving noise coefficient matrix G_j (m by n) and driving noise covariance matrix P_j (n by n positive definite) are all known. (See eq. (2.1).)

If required (usually by a fixed procedure), $E(X_{ij})$ (m by 1) and $Cov(X_{ij})$ (m by m positive definite) will be known, but not dependent on i.

Similarly, for non-geolocation attribute data, given are:

$$B_{\gamma jk} = \begin{pmatrix} B_{\gamma jk1} \\ \vdots \\ B_{\gamma jka_{\gamma jk}} \end{pmatrix} ,$$

$$B_{\gamma jk\eta} = (0, .., 0, 1, 0, .., 0) ,$$

where the 1 occurs in the $b_{\gamma jk\eta}$ th position, $1 \leq \eta \leq a_{\gamma jk}$,

where $1 \leq b_{\gamma jk1} < b_{\gamma jk2} < .. < b_{\gamma jka_{\gamma jk}} \leq b$ and

$$T_{\gamma jk} = \{b_{\gamma jk1}, b_{\gamma jk2}, \ldots, b_{\gamma jka_{\gamma jk}}\} .$$

The corresponding measurement error covariance matrix is

$$R_{\gamma jk} = \begin{pmatrix} r_{b_{\gamma jk1}} & & 0 \\ & \ddots & \\ 0 & & r_{b_{\gamma jk;a_{\gamma jk}}} \end{pmatrix} ,$$

($a_{\gamma jk}$ by $a_{\gamma jk}$ positive definite)

$R_{\gamma jk}$ is a submatrix of fixed by by b positive definite covariance matrix R

$$R = \begin{pmatrix} r_1 & & 0 \\ & \ddots & \\ 0 & & r_b \end{pmatrix} ,$$

which is predetermined.

If normal approximations are not used for the non-geolocation target attributes:

The discrete probability function $pr(Y_{\gamma jk}|H_i)$ is assumed known for all outcomes $Y_{\gamma jk}$ ($a_{\gamma jk}$ by 1) and H_i (b by 1), given selection set $T_{\gamma jk}$. The H_i's are ε C and the Y_{jk}'s are subvectors of vectors ε C.

Quantities Known at t_j , Only After Partitioning $Q^{(j)}$ is Carried Out

$Q^{(j)} = \{Q_0^{(j)}, Q_1^{(j)}, \dots\}$ known, where $Q_i^{(j)} = Z_i^{(j)}$.

f_{jk} the number of false alarms for sensor system k at t_j is known and index i representing target i is now known; all relative to $Q^{(j)}$.

Following rearrangements,

$$Z_j = (Z_{ijk}) \quad \begin{matrix} i=0,1,2,\dots \\ k=i,\dots,q \\ \text{for } Z_{ijk} \neq \phi \end{matrix}$$

$$Z_{ijk} = \begin{pmatrix} \underline{Z}_{ijk} \\ Y_{ijk} \end{pmatrix} = \begin{pmatrix} \underline{Z}'_{\gamma jk} \\ Z''_{\gamma jk} \end{pmatrix} \quad ,$$

for some γ , $1 \leq \gamma \leq m_{jk}$.

$\underline{Z}_{ijk}$ is r_{jk} by 1 geolocation data vector assigned to track set i .

Y_{ijk} is a_{ijk} by 1 geolocation data target attribute vector assigned to track i .

$\underline{Z}_{ijk}$ and/or Y_{ijk} may be vacuous; $Y_{0jk} = \phi$ always.

Thus, T_{ijk}, B_{ijk}, R_{ijk} become determined; similarly for $pr(Y_{ijk}|H_i)$.

The total data $Z^{(j)}$ up to t_j can be broken up two different ways:

$$Z^{(j)} = \{Z_+^{(j)}, Z_0^{(j)}, Y^{(j)}\}$$

$$= \{Z_i^{(j)} | i = 0, 1, 2, \ldots\}$$

where

$$Z_+^{(j)} = \{Z_i^{(j)} | i = 1, 2, \ldots\} \quad ,$$

$$Z_i^{(j)} = \{Z_{i\alpha k} | 0 \leq \alpha \leq j; \; 1 \leq k \leq q, \; Z_{i\alpha k} \neq \phi\}$$

= geolocation data for track set i up to t_j ,

$$Z_0^{(j)} = \{Z_{i\alpha k\omega} | 0 \leq \alpha \leq j \; ; \; 1 \leq k \leq q, \; 1 \leq \omega \leq f_{\alpha k} \; , \; Z_{i\alpha k} \neq \phi\}$$

= false alarm data up to t_j ,

$$Y^{(j)} = \{Y_i^{(j)} | i = 1, 2, ..\} \quad ,$$

$$Y_i^{(j)} = \{Y_{i\alpha k} | \; 0 \leq \alpha \leq j \; ; \; 1 \leq k \leq q, \; Y_{i\alpha k} \neq \phi\}$$

= non-geolocation attribute data for track set i up to t_j .

In turn, these determine sets C_{j1}, C_{j2}, C_{j3} C_{j4}, $\tilde{C}_{j1}$, $\tilde{C}_{j2}$ $\tilde{C}_{j3}$ $\tilde{C}_{j4}$ (the last four are not required, if the Kalman filter-normal approximation procedure is not used for the non-geolocation target attributes). (See Flow boxes 7 - 10 for definitions.)

Note also the notation:

$$\mathbf{Z}_{ij} = (Z_{ijk})_{\substack{1 \leq k \leq q \\ \text{such that} \\ Z_{ijk} \neq \phi}} \quad ; \text{similarly for } Y_{ij} .$$